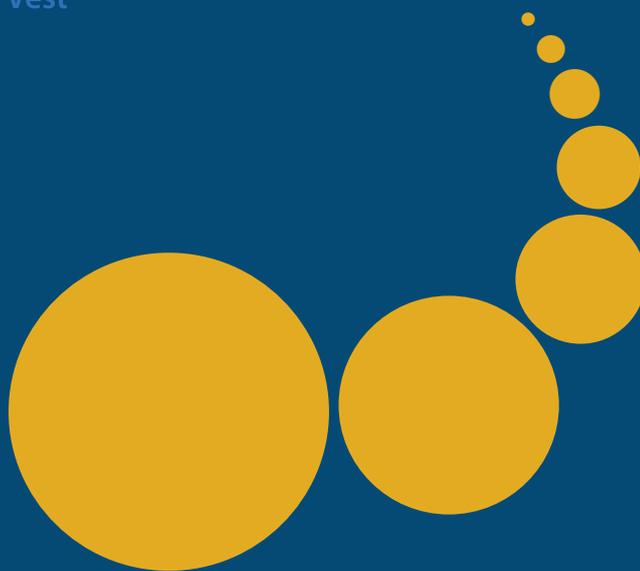


# Scalable Computing: Practice and Experience

---

Scientific International Journal  
for Parallel and Distributed Computing

ISSN: 1895-1767



Volume 25(2)

March 2024

---

EDITOR-IN-CHIEF

**Dana Petcu**

West University of Timisoara, Romania

SENIOR EDITOR

**Marcin Paprzycki**

Systems Research Institute of the Polish Academy of Sciences, Poland

EXECUTIVE EDITOR

**Katarzyna Wasielewska-Michniewska**

Systems Research Institute of the Polish Academy of Sciences, Poland

TECHNICAL EDITOR

**Silviu Panica**

Institute e-Austria Timisoara, Romania

EDITORIAL BOARD

**Peter Arbenz**, Swiss Federal Institute of Technology,

**Giacomo Cabri**, University of Modena and Reggio Emilia,

**Philip Church**, Deakin University,

**Frederic Desprez**, INRIA Grenoble Rhône-Alpes and LIG laboratory,

**Yakov Fet**, Novosibirsk Computing Center,

**Giancarlo Fortino**, University of Calabria,

**Gianluca Frasca-Caccia**, University of Salerno,

**Fernando Gonzalez**, Florida Gulf Coast University,

**Dalvan Griebler**, Pontifical Catholic University of Rio Grande do Sul,

**Frederic Loulergue**, University of Orleans,

**Svetozar Margenov**, Institute for Parallel Processing and Bulgarian Academy of Science,

**Fabrizio Marozzo**, University of Calabria,

**Gabriele Mencagli**, University of Pisa,

**Viorel Negru**, West University of Timisoara,

**Wiesław Pawłowski**, University of Gdańsk,

**Shahram Rahimi**, Mississippi State University,

**Wilson Rivera-Gallego**, University of Puerto Rico,

---

SUBSCRIPTION INFORMATION: please visit <http://www.scpe.org>

# Scalable Computing: Practice and Experience

Volume 25, Number 2, March 2024

---

## TABLE OF CONTENTS

PAPERS IN THE SPECIAL ISSUE ON SCALABILITY AND SUSTAINABILITY IN DISTRIBUTED SENSOR NETWORKS:

<b>Improving the Efficiency and Reliability of Renewable Energy Systems: A Study of Parallel and Distributed Architectures for Integrated Wind and Solar Power Generation</b>	<b>637</b>
<i>Xing Chen, Dingguo Huang, Qingchun Ren, Yong Yang, Ye Yuan</i>	
<b>Implementation of Rules and Routines in Physical Education Teaching and Learning in China</b>	<b>646</b>
<i>Huimin Zhang</i>	
<b>Research on the Design of Integrated Energy Management and Optimization Control Systems for Novel Power Systems</b>	<b>653</b>
<i>Zhiqian Yang, Xianyou Wu, Qihua Chen, Aisikaer, Liangnian Lv, Lei Wu</i>	
<b>Stability Study of New Power System Based on Multi-Intelligent Body Collaboration</b>	<b>661</b>
<i>Xianyou Wu, Zhiqian Yang, Xin Du, Liangnian Lv, Aisikaer, Yanchen Yang</i>	
<b>Study on Grid-Connected Power Quality Improvement of Wind Farms Based on Repetitive Controller</b>	<b>668</b>
<i>Minjie Zhu, Liangnian Lv, Xubo Le, Aisikaer, Haibo Li, Yucheng Gao</i>	
<b>Smart Farming Using the Big Data-Driven Approach for Sustainable Agriculture with IOT – Deep Learning Techniques</b>	<b>675</b>
<i>Buyu Wang</i>	
<b>Scalable Solutions for Wind and Solar Distributed Generation: A Study of Parallel Algorithms in a Smart Grid Environment</b>	<b>683</b>
<i>Jianzhong Li, Feiping Yang, Yu Chen, Li Ao, Wei Xiao</i>	
<b>Physical Education Teaching Quality Evaluation Method Using Mobile Edge Computing in the Online and Offline Environment</b>	<b>692</b>
<i>Huimin Zhang</i>	

<b>Stability Study of Grid-Connected Power System for Wind Farms Considering Power Control</b>	<b>700</b>
<i>Liangnian LV, Zhong Fang, Zhiqian Yang, Aisikaer, Lei Wu, Yanchen Yang</i>	
<b>Research on Speech Communication Enhancement of English Web-based Learning Platform based on Human-computer Intelligent Interaction</b>	<b>709</b>
<i>Yufang Gu</i>	
<b>Optimization Study of Grid Access for Wind Power System Considering Energy Storage</b>	<b>721</b>
<i>Xubo Le, Qiuhua Chen, Minjie Zhu, Yucheng Gao, Bingye Zhang, Aisikaer</i>	
<b>Evaluation of Monitoring Technologies and Methods for Micro Plastics in Water as Novel Pollutants: The Exploration of Accurate Quantitative Analysis and Efficient Screening</b>	<b>729</b>
<i>Ke Hu, Dongdong Li, Xiaolei Cui, Donghua Hu, Junliang Chen, Shaopeng Zhuan, Hao Chang, Yaping Zhang, Tingting An, Juqin Zhang</i>	
<b>A Method for Specifying Yoga Poses Based on Deep Learning, Utilizing OpenCV and Media Pipe Technologies</b>	<b>739</b>
<i>Anuradha T, N. Krishnamoorthy, C S Pavan Kumar, L V Narasimha Prasad, Chunduru Anilkumar, Usha Moorthy</i>	
 <b>PAPERS IN THE SPECIAL ISSUE ON NEXT GENERATION PERVASIVE RECONFIGURABLE COMPUTING FOR HIGH PERFORMANCE REAL TIME APPLICATIONS:</b>	
<b>Interface Control and Status Monitoring of Electronic Information Equipment based on Nonlinear Data Encryption</b>	<b>751</b>
<i>Min Yan, Hua Zhang</i>	
<b>Detection and Prevention of Cyber Defense Attacks using Machine Learning Algorithms</b>	<b>760</b>
<i>Yongqiang Shang</i>	
<b>Security and Privacy of 6G Wireless Communication using Fog Computing and Multi-Access Edge Computing</b>	<b>770</b>
<i>Ting Xu, Ning Wang, Qian Pang, Xiqing Zhao</i>	
<b>Sustainable Development in Medical Applications Using Neural Network Architecture</b>	<b>782</b>
<i>Shuyi Jiang</i>	
<b>Group Intelligent City Mobile Communication Network's Control Strategy based on Cellular Internet of Things</b>	<b>792</b>
<i>Jiazheng Wei</i>	

PAPERS IN THE SPECIAL ISSUE ON MACHINE LEARNING AND BLOCKCHAIN BASED SOLUTION FOR PRIVACY AND ACCESS CONTROL IN IOT:

**Methodology for Developing an IoT-based Parking Space Counter System using XNO** 800  
*Sujanavan Tiruvayipati, Ramadevi Yellasiri, Vikram Narayandas, Archana Maruthavanan, Anupama Meduri*

**A Study of Blockchain and Machine Learning-Enabled IoT Security in Time-Delayed Neural Network Vocal Pattern Recognition to Improve Web-Based Vocal Teaching** 812  
*Kaiyi Long*

PAPERS IN THE SPECIAL ISSUE ON DEEP LEARNING-BASED ADVANCED RESEARCH TRENDS IN SCALABLE COMPUTING:

**Application of Genetic Algorithm in Optimization Simulation of Industrial Waste Land Reuse** 824  
*Peng Bai, Yunan Zhao, Junjia Chang*

**Space Layout Simulation of Assembled Nanoarchitecture Based on Improved Particle Swarm Optimization** 832  
*Huan Huang*

**Optimization of Nonlinear Convolutional Neural Networks based on Improved Chameleon Group Algorithm** 840  
*Qingtao Zhang*

**Channel Estimation of Urban 5G Communication System based on Improved Particle Swarm Optimization Algorithm** 848  
*Xigang Xia, Bo Yang, Zhiyu Liu*

**Multi-source and Multi-level Coordination of Energy Internet under V2G based on Particle Swarm Optimization Algorithm** 857  
*Jian Xu, Yunyan Chang, Xiaoming Sun*

**Numerical Simulation and Optimal Control of Composite Nonlinear Mechanical Parts Casting Process** 867  
*Huan Li, Peng Wang*

**Conservation Design of Industrial Heritage based on Nonlinear GA Optimization Algorithm and Three-dimensional Reconstruction** 874  
*Yunan Zhao, Peng Bai*

**Lightweight Saliency Target Intelligent Detection based on Multi-scale Feature Adaptive Fusion** 883  
*Muqing Zhu*

<b>Optimization of Radio Energy Transmission System Efficiency Based on Genetic Algorithm</b>	<b>891</b>
<i>Ruijuan Du</i>	
<b>Intelligent Detection and Analysis of Software Vulnerabilities based on Encryption Algorithms and Feature Extraction</b>	<b>900</b>
<i>Heng Li, Xinqiang Li, Hongchang Wei</i>	
<b>Application of Control Algorithm in the Design of Automatic Crimping Device for Connecting Pipe and Ground Wire</b>	<b>909</b>
<i>Congbing Sheng, Peng Xing, Xiuzhong Cai, Zheng Shao</i>	
<b>Linear Anti-interference Algorithm for Digital Signal Transmission in Fiber Optic Communication Networks based on Link Analysis</b>	<b>920</b>
<i>Jing Wu, Cheng Jin, Ziwu Wang</i>	
<b>Network Traffic Monitoring and Real-time Risk Warning based on Static Baseline Algorithm</b>	<b>928</b>
<i>Zhaoli Wu, Junwei Liu</i>	
<b>Application of Improved PSO and BP Hybrid Optimization Algorithm in Electrical Automation Intelligent Control System</b>	<b>938</b>
<i>Lijing Li, Xiaojian Wang, Mei Yang</i>	
<b>Design of Computer Information Management System Based on Machine Learning Algorithms</b>	<b>944</b>
<i>Yan Li</i>	
<b>Automatic Control of Low Voltage Load in Power Systems Based on Deep Learning</b>	<b>952</b>
<i>Yaohui Sun, Hongyu Zhang, Haolin Li, Shu Wang, Chunhai Li</i>	
<b>Target Image Processing Based on Super-resolution Reconstruction and Deep Machine Learning Algorithm</b>	<b>961</b>
<i>Yang Lin, Ping Zhang, He Zhang, Guoping Song</i>	
 <b>PAPERS IN THE SPECIAL ISSUE ON SCALABLE COMPUTING IN ONLINE AND BLENDED LEARNING ENVIRONMENTS: CHALLENGES AND SOLUTIONS:</b>	
<b>Text Summarization for Online and Blended Learning</b>	<b>987</b>
<i>Mahira Kirmani, Gagandeep Kaur</i>	
<b>The Effects of Integrated Feedback based on AWE on English Writing of Chinese EFL Learners</b>	<b>987</b>
<i>Mei Liu, Changzhong Shao</i>	

PAPERS IN THE SPECIAL ISSUE ON SCALABLE DEW COMPUTING FOR FUTURE GENERATION  
IoT SYSTEMS:

**An Elixir for Blockchain Scalability with Channel based Clustered Sharding** 997

*V. Vinoth Kumar, U. Padmavathi, C. Prasanna Ranjith, Balaji J, C.N.S.Vinoth Kumar*

**Feature Extraction and Classification of Gray-Scale Images of Brain Tumor using Deep Learning** 1005

*Kondra Pranitha, Naresh Vurukonda*

**A Class Specific Feature Selection Method for Improving the Performance of Text Classification** 1018

*Venkatesh V, Sharan S B, Mahalaxmy S, S Monisha, Ashick Sanjey D S, Ashokkumar P*

**Breast Cancer Image Classification based on Adaptive Interpolation Approach Using Clinical Dataset** 1029

*Susmitha Uddaraju, A Koushlya, I Hemalatha, M Maragatharajan, Balasubramanian C and L Sathish Kumar*

**SecureSense: Enhancing Person Verification through Multimodal Biometrics for Robust Authentication** 1040

*Samatha J, G. Madhavi*

**Radiogenomics in Oncology: A Comprehensive Study of Various Oncological Disorders** 1055

*Sowmya V L, Bharathi Malakreddy A, Santhi Natarajan*

**Sensor based Dance Coherent Action Generation Model using Deep Learning Framework** 1073

*Hanzhen Jiang, Yingdong Yan*

**IoT based Dance Movement Recognition Model based on Deep Learning Framework** 1091

*Zhen Ji, Yaonong Tian*

PAPERS IN THE SPECIAL ISSUE ON GRAPH POWERED BIG AEROSPACE DATA PROCESSING:

**Research on Network Security Situation Awareness Technology Based on Security Intelligent Monitoring Technology** 1107

*Bingyu Yang*

**A Distributed System Fault Diagnosis System based on Machine Learning** 1117

*Yixiao Wang*

**Intelligent Navigation System based on Big Data Traffic System** **1124**

*Xiu Zhang, Jian Kang, Haicun Yu*

PAPERS IN THE SPECIAL ISSUE ON SOFT COMPUTING AND ARTIFICIAL INTELLIGENCE FOR WIRE/WIRELESS HUMAN-MACHINE INTERFACE:

**VisiSense: A Comprehensive IOT-based Assistive Technology System for Enhanced Navigation Support for the Visually Impaired** **1134**

*Bhasha Pydala, T. Pavan Kumar, K. Khaja Baseer*

**Optimum Batch Scheduling Model for Quality Aware Delay Sensitive Data Transmission over Fog Enabled IOT Network** **1152**

*Narayana Potu, Chandrashekar Jatoth, Premchand Parvataneni*

**Trajectory Interception Classification for Prediction of Collision Scope between Moving Objects** **1167**

*B. Uma Mahesh Babu, K. Giri Babu, B. T. Krishna*

**Security Enabled New Term Weight Measure Technique with Data Driven for Next Generation Mobile Computing Networks** **1191**

*Anil Kumar Budati, Shayla Islam, Shaik Mohammad Rafee, Chengamma Chitteti, T. Lakshmi Narayana*

**Optimizing Multichannel Path Scheduling in Cognitive Radio Ad Hoc Networks using Differential Evolution** **1199**

*Dasari Ramesh, N. Venkatram*

**Iterative Ensemble Learning over High Dimensional Data for Sentiment Analysis** **1219**

*V R N S S V Saileela P, N. Naga Malleswara Rao*

**Optimal Usage of Resources through Quality Aware Scheduling in Containers based Cloud Computing Environment** **1235**

*S.A. Poojitha, K. Ravindranath*

**Enhanced Feature Optimization for Multiclass Intrusion Detection in IOT Fog Computing Environments** **1246**

*Sudarshan S. Sonawane*

**Quality Enhancement with Frame-wise DLCNN using High Efficiency Video Coding in 5G Networks** **1264**

*Vijaya Saradhi Dommeti, M. Dharani, K. Shasidhar, Y Dasaratha Rami Reddy, T. Venkatakrisna Moorthy*

RESEARCH PAPERS:

**Software Effort Estimation using Machine Learning Algorithms** **1276**

*Kruti Lavingia, Raj Patel, Vivek Patel, Ami Lavingia*

---

© SCPE, Timișoara 2024





## IMPROVING THE EFFICIENCY AND RELIABILITY OF RENEWABLE ENERGY SYSTEMS: A STUDY OF PARALLEL AND DISTRIBUTED ARCHITECTURES FOR INTEGRATED WIND AND SOLAR POWER GENERATION

XING CHEN, DINGGUO HUANG, QINGCHUN REN, YONG YANG, AND YE YUAN\*

**Abstract.** The implications of relevant sustainable practices have reflected the scholastic features to improve the environmental resources. The study highlights the importance of conservation of environment and power system in search of proven solutions to improve the penetration level. The need for flexibility has signified the special characteristics that are conventional in increasing the integrity of renewable resources. The ideologies have global trends have integrated the cost of affectivity with the growing applications of power projects. The architecture of wind and solar energy has touched successful benchmarks with respect to the real world implications. The conceptual practices help in initializing the practices towards biomass as well as determining the impact of renewable energy on wind and solar power energy in a significant manner. In addition to that, the applications of solar or photovoltaic cell have been mentioned in the study which has greater significance. The ideas based on the emission of greenhouse gases have been evaluated in the study that shows the after effects as well. The use of passive solar energy and active solar energy has clearly discussed the concept of sustainability and the process of administering towards various climatic conditions. Lastly, the impact of renewable resources on social, environmental, technical and economic aspects has verified the relevant practice of sustainability.

**Key words:** Sustainability, renewable sources, passive solar energy, active solar energy, reliability, greenhouse gases

**1. Introduction.** Energy efficiency simply refers to reducing excessive energy use to perform continuous tasks. A decrease in energy use brings up a variety of benefits, including knowing the greenhouse gas emission and reducing the cost value at the economic and household levels. While renewable technologies are readily helpful in accomplishing the objectives and are a way of reducing fossil fuels. This study explores how to make renewable energy systems more reliable and effective. The project investigates the potential of parallel and distributed architectures with a focus on integrated wind and solar power generation. The study seeks to maximise the effectiveness and dependability of renewable energy systems by examining these methods, offering significant insights into the world of sustainable energy.

The motivation behind this abstract lies in addressing the critical need for sustainable practices in the context of environmental conservation and power systems. It recognizes the pressing global concern of environmental degradation and the importance of seeking viable solutions to increase the adoption of sustainable energy sources. This study aims to shed light on the significance of preserving the environment while enhancing the penetration and reliability of renewable resources.

The primary aim of this research is to explore and emphasize the special characteristics of renewable energy sources, such as wind and solar power, as key components in achieving a sustainable future. By delving into the practical implications of these technologies, the study seeks to establish real-world benchmarks and applications that can contribute to a cleaner and more sustainable energy landscape. Furthermore, the research aspires to examine the impact of renewable energy sources, particularly in the realms of wind and solar power, with a specific focus on biomass. It aims to comprehensively assess their contributions and implications, thereby providing insights into the significance of renewable energy in both local and global contexts.

### 2. Objectives.

1. To evaluate the importance of renewable energy in improving the reliability of the ecosystem
2. To determine the impact of renewable energy on wind and solar power energy
3. To understand the distributed architecture for the supervision of energy management

---

\*Chongqing Qingdian New Energy Development Co. Ltd., Chongqing,401121, China ([dingguohuang21@outlook.com](mailto:dingguohuang21@outlook.com))

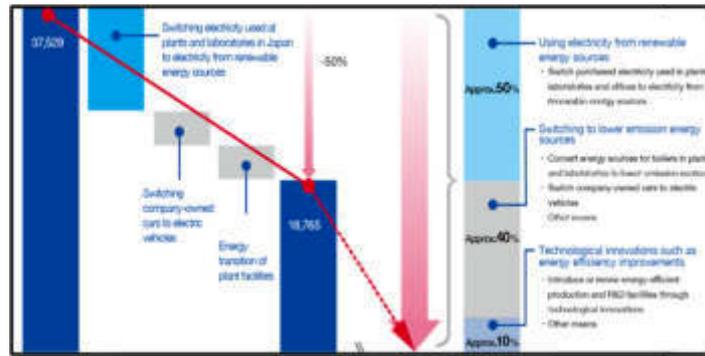


Fig. 3.1: Effectiveness of renewable energy on the ecosystem

#### 4. To analyze the challenges faced during the adoption of renewable energy

**3. Methodology.** The analysis process includes a wide spectrum of knowledge based on the analytical approaches in composing the literature portion [26]. Secondary data has been identified as relevant in terms of analyzing the entire study which comprises statistical and numerical information. Secondary data imposes a wider scope in analyzing the importance of renewable energy resources proclaiming to generate ideas on wind and solar energy [1]. The phenomenon is well recognized by the qualitative data having a valuable outcome in elaborating the section on the efficiency and reliability of renewable energy resources.

**3.1. Evaluating the importance of renewable energy in improving the reliability of the ecosystem.** Renewable energy has no such exception to the rule of thumb and has its own ways of implementing the improvement of solar and wind energy. Renewable energy is good for the planet as it helps in reducing excessive pollution and it brings social cohesion as well [2]. The pollution caused by the combustion of greenhouse gases is controlled by the implication of renewable energy.

*Renewable energy has zero emission of greenhouse gases.* The combustion of fossil fuels is significant in contributing the global warming. Most renewable resources are useful in terms of reducing global complexities by considering the full life cycle of technologies [3]. The increase in fossil fuels worldwide has resulted in excessive forms of global warming. Such issues can be controlled by the application of renewable practices.

*Renewable energy is cost-effective.* The use of renewable energy is beneficial in many ways mostly it is highly cost-effective in nature. Since renewable resources have been a great need to fulfill the geopolitical crisis, the upheavals of energy resources are often considered calculative in decentralizing climate change [4]. As a result of this, the recollection of technology and an aversive use of renewable projects can lead to effective results.

*Renewable energy is accessible in nature.* Renewable energy is highly accessible to various parts of the world and it is more independent from remote success. Especially for cities, an energy based on distributed and decentralized knowledge is significant in avoiding weather-related fluctuations [5]. Business and trade industries are highly benefitted by the use of urban energy to expand the rate of accessibility in suburban and peri-urban regions.

**3.2. Impact of renewable energy on wind and solar power energy.** Renewable energy is derived from natural sources that aim in replenishing the higher rates of pollutants. Wind and solar energy does not generate atmospheric thermal pollution or contaminants. Rather, they reduce atmospheric emissions which mainly accumulate greenhouse gases, carbon dioxide (CO<sub>2</sub>), nitrogen dioxide (NO<sub>x</sub>), and the particulates of sulfur dioxide (SO<sub>2</sub>) [6].

*Active solar heating.* Active heating enables the process of collecting and administering the distributed approaches that create a source of solar energy that enables the transfer of carrier fluids. It functions with the help of solar energy by transferring the hot fluid which can be either air or liquid [7]. Active solar heating

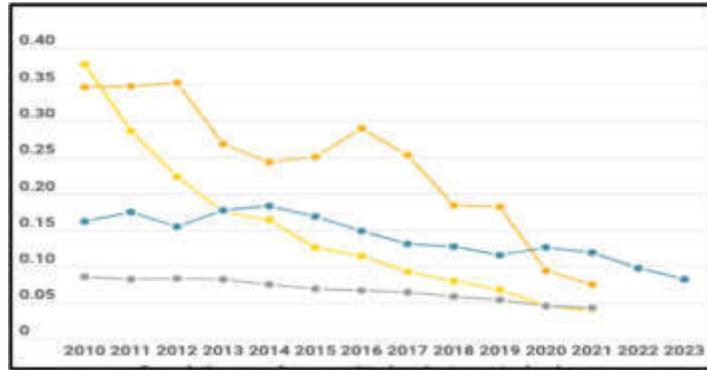


Fig. 3.2: Effective courses of cost between the years 2020 to 2023 (Influenced by [4])

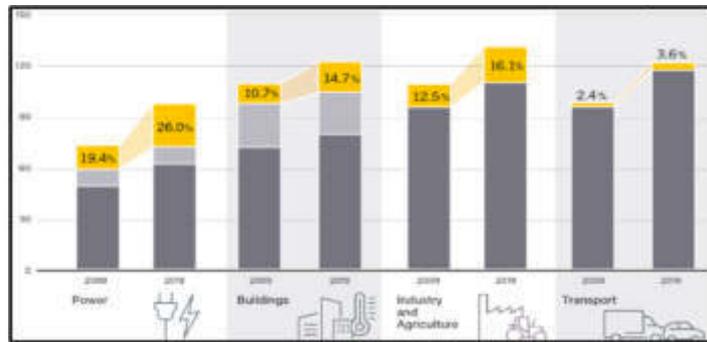


Fig. 3.3: Rate of accessibility in renewable energy

reduces energy consumption having diverse applications. It hardly takes any maintenance cost and manages to work in various climatic conditions.

*Passive solar heating.* Passive solar heating is a necessary mechanism that is used to protect buildings from excessive ultraviolet sun rays. It is a well-designed system that reduces heating and increases the cooling effect of the buildings. The goal of passive solar heating is to absorb the excess heat and maintain a comfortable room temperature [8].

In the structure shown in figure 3.5, it has been illustrated that passive solar technologies can be used both in a direct and indirect manner to gain spacing for thermal mass based on thermo siphon. As a result, passive solar energy minimizes the emission of carbon into the environment which is a much more conventional process in the heating systems in a significant manner [9].

*Application of solar or photovoltaic cell.* Solar electricity generates the use of electrical powers that helps in forming the process of the semiconductor devices that are closely associated with computer chip programming. Photovoltaic cells are useful in generating electrical energies that mainly comprise electrons connected to the utility of sound alternatives [10].

Figure 3.6 illustrates the semiconductor material is the core tool that acts as a protective layer and stops the ultraviolet rays from penetrating into the surface. Through solar cells, environmental sustainability can be achieved which makes it ideal for the ideal mechanism of operations and maintenance of cost cells [11].

### 3.3. Understanding the distributed architecture for the supervision of energy management .

Investing in the electricity distribution generated from renewable energy develops various energy resources such as solar, wind, biomass, and ocean waves. With the development in technology, it can be administered that affordable power distribution is competent for energy management [12]. It helps in supervising the elements

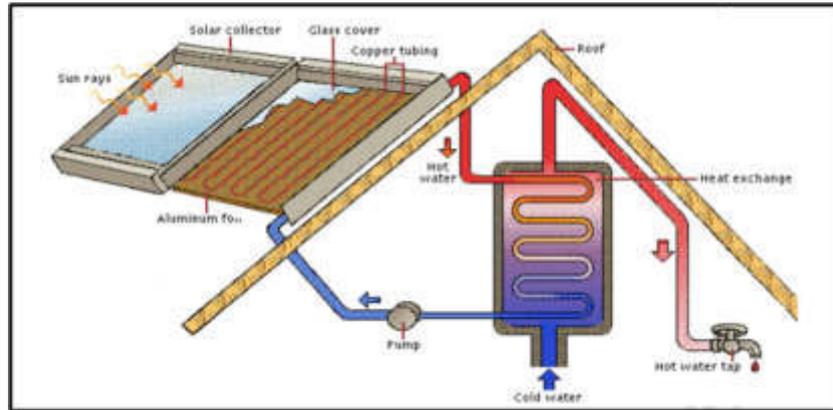


Fig. 3.4: Active solar heating

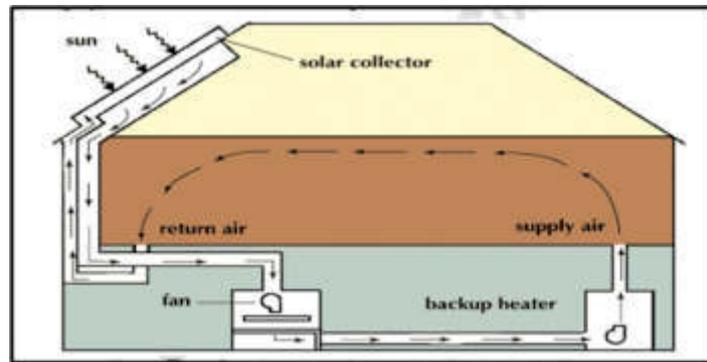


Fig. 3.5: Passive solar heating

that improve supply management practices. Similarly the switching of the system from a central controller to inter-communicative tools has bridged the gap made by the challenges. In addition to that, the power-generating paradigm ensures the safety and services towards autonomous sustainability [13]. The supervision of energy has uplifted sustainability practices through renewable resources.

*Serial converter connection.* The isolation process has led to the emergence of DC-DC converters in terms of generating the basics of solar and wind energy. The connectivity gets stronger with a series of high-voltage output stabilizers in order to reach the desired output [14].

Figure 3.7 illustrates the current practices involving renewable practices in accordance with outreaching the converter outputs. The high resolution has enriched the goals of the breakdown voltage of the output that has been programmed simultaneously [15]. The adaptation of serial converter connection includes various advantages:

1. It uses less number of wires
2. It associates the path of connectivity between receiving and transmitting devices
3. It supports long-distance communication that
4. It encourages the usage of renewable resource practices

*Cascading converter connection.* The gain in the rate of voltage is theoretically measured that is inductive in nature accumulating the robust applications of electricity distribution. After the introduction of renewable energy, the relevant approaches have converted the regular use of energy transfer into the management process [16].

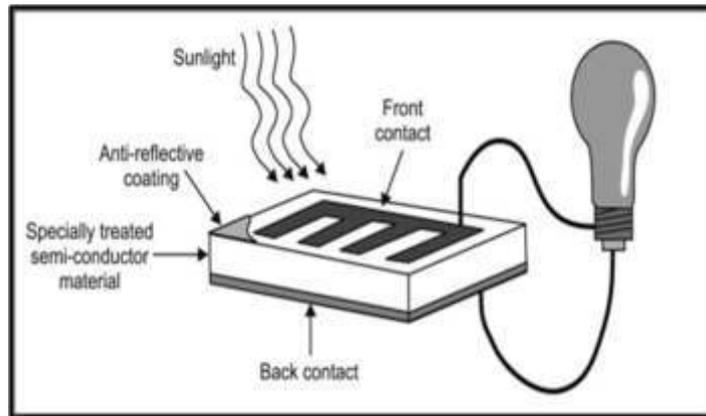


Fig. 3.6: Solar or Photovoltaic cell

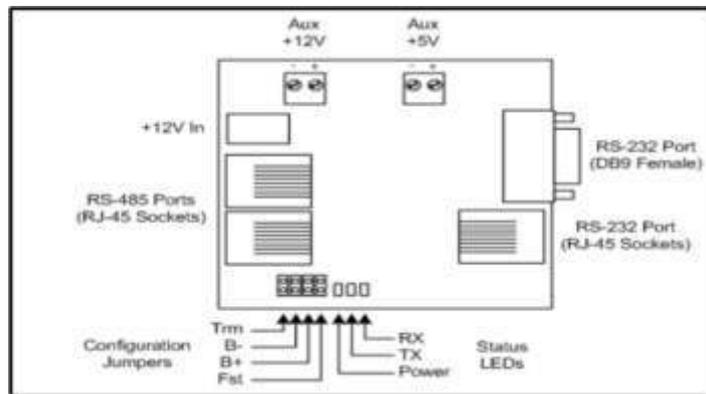


Fig. 3.7: Serial converter connection

The above system is based on various modalities with respect to the initiation of cascading arrangements [17]. There are various applications of the cascading converter connection which mainly include:

1. Predicting the system performance under abnormal conditions
2. A static and comprehensive way towards dynamic performance
3. Robust source of connectivity between the individual PV modules
4. Providing alternative paths to promote renewable practices

*Connecting converters in parallel form.* In the augmentation of the power supply, the requirement for connectivity has often led to the optimization of the distribution system. This distribution is a mode of inquiring about the distribution of practices with respect to the homogenous distribution of power supply [18].

The above structure shown in figure 3.9 possesses the exact features that are relevant to the dispersion of connectivity. This phenomenon is an ultimatum that has cultivated the practice of stressing semiconductor resources [19]. In comparison to homogenous distribution, the power project has various relevant implications for sustainability.

1. The mechanism technically generates output powering
2. It manages the issues of overload by improving the connectivity sources
3. It is intended to maximize the load-sharing practices toward redundant operations
4. The learning parameters are adjusted and converted to the algorithm

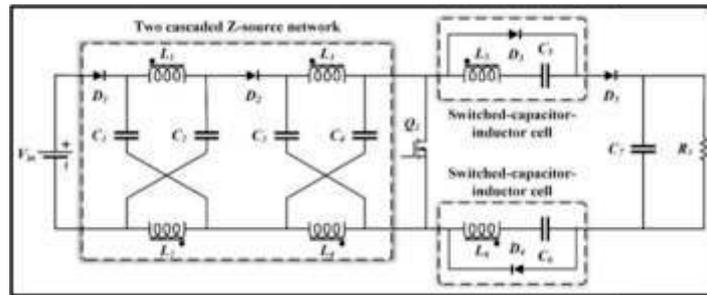


Fig. 3.8: Cascading converter connection

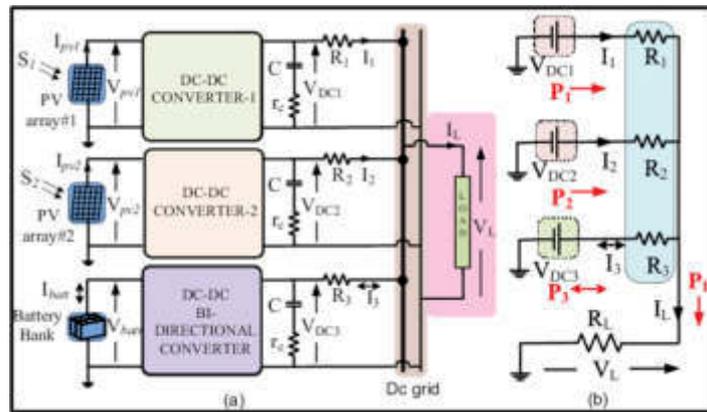


Fig. 3.9: Connecting converters in parallel form

**3.4. Challenges faced during the adoption of renewable energy.** Renewable technology has become better with time progress and it has access to all the resources that are responsible to contribute nearly 20% to the consumption of global energy. Wind and solar energy are bringing about clearer and more innovative ways in order to capture the cost of energy [20]. The challenges have induced pressure on the environment thereby disrupting the practice of offering renewable resources. It can be measured that non-renewable resources can be found in certain parts of the environment. Such practices are a predicament of ideas that are approachable to the course of limited amounts of sources [21].

*High initial costs of installation.* Carbon emissions are the main reason for global warming and it has been observed that in order to increase the adaptation renewable energy needs proper installation. Solar as well as wind energy are the cheapest ones and are used widely [22].

However, the figure 3.10 highlights the installation process which is abnormally high and is around 21% between the years 2021 to 2022 which is \$2,000 per kilowatt. There is a wide margin in the installation of fossil fuel plants. Moreover, due to high-power storage, the installation becomes even more difficult [23].

*Lack of policies and subsidies.* The lack of policies and subsidies is a hindrance to the installation of renewable energy. In addition to that, political stigma, corporation lobby, and inherent dependencies have been challenging during the installation process [24].

The figure 3.11 establishes the dangers of climatic depletion that have started to affect the lives of people which eventually lead to the loss of the ecosystem. In addition to that, the high property taxes on fossil fuels have been a threat to the local community in order to establish sustainable energy [25].

**4. Results.** The above systematic diagram shown in figure 4.1 illustrates that an increased rate of sustainability can lead to various effective outcomes. The various impacts of renewable energy can be demonstrated

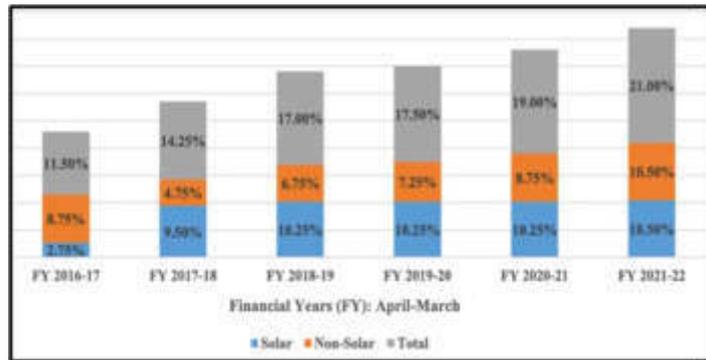


Fig. 3.10: Cost factor in the installation of renewable energy

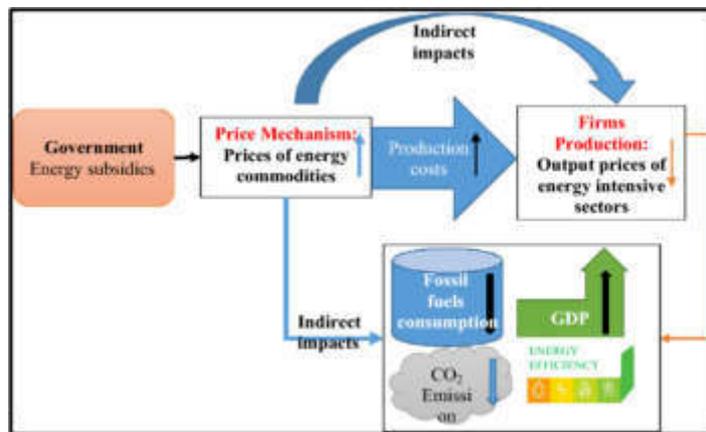


Fig. 3.11: Policy investments in the installation of sustainability

in the form of a comprehensive understanding of the global environment. The varied elements can be discussed as follows:

*Social and environmental impacts.* Renewable energy has played a crucial role in foregoing the practice of reliable and sustainable resources. This has reduced the emission of harmful gases producing greenhouse gases which is less than that is between 90-99% as compared to coal fire plants of 70-90% significantly [27]. Additionally, social impacts comprise the elimination of poverty and change in climate mitigation.

*Technical aspects.* The evolution of technology has clearly exceeded all expectations on the verge of sustainable practices. Technical synergies have played a large role in the primary supply of energy [28]. Renewable practices have raised the range of connectivity measuring technical support to improve the electricity demands.

*Economic factors.* Renewable practices have raised human well-being and have developed the overall GDP growth rate. By doubling up the share, it has increased the global rate by 1.1% significantly [29]. This is related to USD 1.3 trillion with respect to the service of fossil fuels.

*Commercialization.* Renewable energy commercialization involves the additional factors that have improved the research knowledge. This has contributed to 19% of the energy consumption towards photovoltaic thermal power stations [30].

**5. Conclusion.** In conclusion, the research presented in this study underscores the critical role of semiconductor materials in protecting against ultraviolet rays and enabling the use of solar cells for achieving environmental sustainability. It highlights the potential of renewable energy sources, including solar, wind,

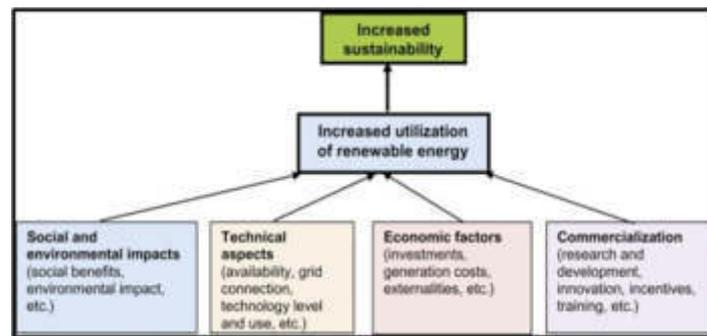


Fig. 4.1: Impact of sustainability

biomass, and ocean waves, in reshaping energy management practices and supporting the transition to cleaner, more sustainable power generation methods. The investigation further delves into the advantages of different converter connection methods, including serial, cascading, and parallel forms, to optimize energy distribution and ensure the efficient operation of renewable energy systems. The challenges faced during the adoption of renewable energy are recognized, particularly the high initial installation costs and the lack of supportive policies and subsidies. These challenges, along with the influence of political and corporate interests, pose significant hurdles in transitioning to cleaner energy sources. Nonetheless, the research underscores the urgency of addressing these challenges to mitigate environmental degradation and ensure a sustainable energy future. The findings suggest that renewable energy can significantly reduce greenhouse gas emissions, alleviate poverty, and drive economic growth, ultimately leading to a more sustainable and environmentally responsible energy landscape.

## REFERENCES

- [1] M. S. ABID, H. J. APON, I. M. NAFI, A. AHMED, AND R. AHSHAN, *Multi-objective architecture for strategic integration of distributed energy resources and battery storage system in microgrids*, *Journal of Energy Storage*, 72 (2023), p. 108276.
- [2] M. S. ALAM, F. S. AL-ISMAIL, A. SALEM, AND M. A. ABIDO, *High-level penetration of renewable energy sources into grid utility: Challenges and solutions*, *IEEE Access*, 8 (2020), pp. 190277–190299.
- [3] M. L. ALGHAYTHI, R. M. O’CONNELL, N. E. ISLAM, M. M. S. KHAN, AND J. M. GUERRERO, *A high step-up interleaved dc-dc converter with voltage multiplier and coupled inductors for renewable energy systems*, *IEEE Access*, 8 (2020), pp. 123165–123174.
- [4] A. S. ALSAFRAN, *A feasibility study of implementing iec 1547 and iec 2030 standards for microgrid in the kingdom of saudi arabia*, *Energies*, 16 (2023), p. 1777.
- [5] S. ASLAM, H. HERODOTOU, S. M. MOHSIN, N. JAVAID, N. ASHRAF, AND S. ASLAM, *A survey on deep learning methods for power load and renewable energy forecasting in smart microgrids*, *Renewable and Sustainable Energy Reviews*, 144 (2021), p. 110992.
- [6] P. R. BANA, K. P. PANDA, R. NAAYAGI, P. SIANO, AND G. PANDA, *Recently developed reduced switch multilevel inverter for renewable energy integration and drives application: topologies, comprehensive analysis and comparative evaluation*, *IEEE access*, 7 (2019), pp. 54888–54909.
- [7] C. BARROWS, A. BLOOM, A. EHLEN, J. IKÄHEIMO, J. JORGENSEN, D. KRISHNAMURTHY, J. LAU, B. MCBENNETT, M. O’CONNELL, E. PRESTON, ET AL., *The iec reliability test system: A proposed 2019 update*, *IEEE Transactions on Power Systems*, 35 (2019), pp. 119–127.
- [8] J. A. DOMINGUEZ, N. HENAO, K. AGBOSSOU, A. PARRADO, J. CAMPILLO, AND S. H. NAGARSHETH, *A stochastic approach to integrating electrical thermal storage in distributed demand response: A case study for nordic communities with wind power generation*, *IEEE Open Journal of Industry Applications*, (2023).
- [9] M. M. GULZAR, A. IQBAL, D. SIBTAIN, AND M. KHALID, *An innovative converterless solar pv control strategy for a grid connected hybrid pv/wind/fuel-cell system coupled with battery energy storage*, *IEEE Access*, 11 (2023), pp. 23245–23259.
- [10] M. HANNAN, M. H. LIPU, P. J. KER, R. BEGUM, V. G. AGELIDIS, AND F. BLAABJERG, *Power electronics contribution to renewable energy conversion addressing emission reduction: Applications, issues, and recommendations*, *Applied energy*, 251 (2019), p. 113404.
- [11] S. IMPRAM, S. V. NESE, AND B. ORAL, *Challenges of renewable energy penetration on power system flexibility: A survey*, *Energy Strategy Reviews*, 31 (2020), p. 100539.
- [12] W. JIANG, X. WANG, H. HUANG, D. ZHANG, AND N. GHADIMI, *Optimal economic scheduling of microgrids considering*

- renewable energy sources based on energy hub model using demand response and improved water wave optimization algorithm*, Journal of Energy Storage, 55 (2022), p. 105311.
- [13] M. KHALEEL, Z. YUSUPOV, M. ELMNIFI, T. ELMENFY, Z. RAJAB, AND M. ELBAR, *Assessing the financial impact and mitigation methods for voltage sag in power grid*, International Journal of Electrical Engineering and Sustainability (IJEES), (2023), pp. 10–26.
- [14] T. KONSTANTINOY AND N. HATZIARGYRIOU, *Day-ahead parametric probabilistic forecasting of wind and solar power generation using bounded probability distributions and hybrid neural networks*, IEEE Transactions on Sustainable Energy, (2023).
- [15] Z. LIU, K. HOU, H. JIA, J. ZHAO, D. WANG, Y. MU, AND L. ZHU, *A lagrange multiplier based state enumeration reliability assessment for power systems with multiple types of loads and renewable generations*, IEEE Transactions on Power Systems, 36 (2020), pp. 3260–3270.
- [16] L. LV, Z. WU, L. ZHANG, B. B. GUPTA, AND Z. TIAN, *An edge-ai based forecasting approach for improving smart microgrid efficiency*, IEEE Transactions on Industrial Informatics, 18 (2022), pp. 7946–7954.
- [17] G. MA, J. LI, AND X.-P. ZHANG, *A review on optimal energy management of multimicrogrid system considering uncertainties*, IEEE Access, 10 (2022), pp. 77081–77098.
- [18] D. E. OCHOA, F. GALARZA-JIMENEZ, F. WILCHES-BERNAL, D. SCHOENWALD, AND J. I. POVEDA, *Control systems for low-inertia power grids: A survey on virtual power plants*, IEEE Access, (2023).
- [19] O. OGUNDAIRO, S. KAMALASADAN, A. R. NAIR, AND M. SMITH, *Oscillation damping of integrated transmission and distribution power grid with renewables based on novel measurement-based optimal controller*, IEEE Transactions on Industry Applications, 58 (2022), pp. 4181–4191.
- [20] S. G. PAUL, A. SAHA, M. S. AREFIN, T. BHUIYAN, A. A. BISWAS, A. W. REZA, N. M. ALOTAIBI, S. A. ALYAMI, AND M. A. MONI, *A comprehensive review of green computing: Past, present, and future research*, IEEE Access, (2023).
- [21] P. ROY, J. HE, T. ZHAO, AND Y. V. SINGH, *Recent advances of wind-solar hybrid renewable energy systems for power generation: A review*, IEEE Open Journal of the Industrial Electronics Society, 3 (2022), pp. 81–104.
- [22] S. SEN AND V. KUMAR, *Distributed adaptive-mpc type optimal pms for pv-battery based isolated microgrid*, IEEE Systems Journal, 17 (2022), pp. 546–557.
- [23] K.-Y. TAI, F. Y.-S. LIN, AND C.-H. HSIAO, *An integrated optimization-based algorithm for energy efficiency and resource allocation in heterogeneous cloud computing centers*, IEEE Access, (2023).
- [24] M. THIRUNAVUKKARASU, Y. SAWLE, AND H. LALA, *A comprehensive review on optimization of hybrid renewable energy systems using various optimization techniques*, Renewable and Sustainable Energy Reviews, 176 (2023), p. 113192.
- [25] M. TOURE, M. CAMARA, A. PAYMAN, AND B. DAKYO, *African renewable energy potentialities review for local weak grids reinforcement study*, in 2023 11th International Conference on Smart Grid (icSmartGrid), IEEE, 2023, pp. 1–9.
- [26] Q. WANG, F. LI, Y. TANG, AND Y. XU, *Integrating model-driven and data-driven methods for power system frequency stability assessment and control*, IEEE Transactions on Power Systems, 34 (2019), pp. 4557–4568.
- [27] Z. WANG, Z. ZHOU, H. ZHANG, G. ZHANG, H. DING, AND A. FAROUK, *AI-based cloud-edge-device collaboration in 6g space-air-ground integrated power iot*, IEEE Wireless Communications, 29 (2022), pp. 16–23.
- [28] L. XIONG, Y. TANG, S. MAO, H. LIU, K. MENG, Z. DONG, AND F. QIAN, *A two-level energy management strategy for multi-microgrid systems with interval prediction and reinforcement learning*, IEEE Transactions on Circuits and Systems I: Regular Papers, 69 (2022), pp. 1788–1799.
- [29] K. Y. YAP, H. H. CHIN, AND J. J. KLEMEŠ, *Future outlook on 6g technology for renewable energy sources (res)*, Renewable and Sustainable Energy Reviews, 167 (2022), p. 112722.
- [30] Z.-P. YUAN, P. LI, Z.-L. LI, AND J. XIA, *Data-driven risk-adjusted robust energy management for microgrids integrating demand response aggregator and renewable energies*, IEEE Transactions on Smart Grid, 14 (2022), pp. 365–377.

*Edited by:* Sathishkumar V E

*Special issue on:* Scalability and Sustainability in Distributed Sensor Networks

*Received:* Aug 23, 2023

*Accepted:* Nov 16, 2023



## IMPLEMENTATION OF RULES AND ROUTINES IN PHYSICAL EDUCATION TEACHING AND LEARNING IN CHINA

HUIMIN ZHANG\*

**Abstract.** Physical education is a significant aspect of the Chinese education system. Moreover, on a regular basis, physical education is an integral part of the Chinese education system. Therefore, the following study looks into different rules and routines for physical education teaching and learning in China. Most of the rules and regulations are formed based on guidelines provided by the Ministry of Education. Therefore, a systematic discussion regarding physical education training in China is conducted in the analysis. The significance of physical education (PE) in the Chinese education system cannot be overstated. It serves a dual purpose by not only promoting physical fitness but also fostering holistic personal development. PE contributes to the physical well-being of students, helping them lead healthier lives, but it also instills essential life skills like teamwork, discipline, and perseverance. In China, the Ministry of Education, as the overarching authority on educational matters, plays a pivotal role in shaping the rules and routines governing physical education. These rules encompass a wide array of aspects related to the curriculum, including the allocation of resources, curriculum design, assessment, and teaching methodologies. By adhering to these regulations, educational institutions across the country can ensure a standardized and comprehensive approach to physical education.

**Key words:** PE education, the benefit of physical education, Resource allocation of PE education in China

**1. Introduction.** For Chinese students, it is essential to follow a regular physical education routine along with their regular studies. Moreover, the education ministry of China has provided strict guidelines regarding the physical education system that helped to formulate the base of regular physical training [1]. Therefore, through a secondary analysis, the following study looks into the implementation of rules and routines in physical education teaching and learning in China. At the same time, the study has discussed the problem of the same. Through a secondary analysis, a coherent discussion related to possible solutions is discussed in the study. Additionally, results are formulated based on the information gathered from the secondary data.

Physical education (PE) is an integral and indispensable component of educational systems worldwide. It encompasses a broad spectrum of activities and teachings designed to promote physical fitness, health, and well-being among students. In addition to fostering physical health, physical education plays a significant role in enhancing cognitive, emotional, and social development. This multifaceted approach to education is particularly significant in China, where it has been woven into the fabric of the education system for many years.

The Chinese education system places immense importance on physical education, recognizing that a healthy body is closely linked to a healthy mind. Beyond the benefits of improved fitness, physical education imparts values such as discipline, teamwork, and perseverance, which are integral to a student's holistic development. This system's commitment to physical education has led to the formulation of a set of rules and routines that govern the teaching and learning of physical education in China. These rules, often dictated by the Ministry of Education, provide a framework for curriculum development, resource allocation, and assessment practices within the realm of PE.

**2. Objectives.** For the development of the empirical analysis and formulation of reliable results following objectives were followed.

1. To discuss the elements impacting the implementation of rules and routines in physical education teaching and learning in China
2. To analyze the impact of physical education and teaching on the Chinese students

---

\*School of Public Teaching and Practice Wuhan Technical College of Communications, Wuhan, 430065, China ([huiminzhang21@outlook.com](mailto:huiminzhang21@outlook.com))

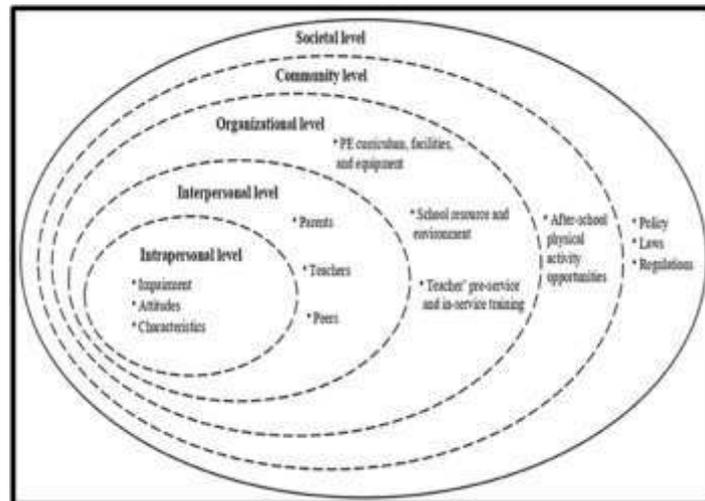


Fig. 1.1: Benefits of physical education in different strata

3. To elaborate on the issues related to the implementation of rules and routines in physical education teaching and learning in China
4. To discuss possible results and solutions in order to improve the physical education teaching and learning for Chinese students

**3. Methodology.** The methodology of an empirical analysis looks into the factors that aided in the development of the analysis. Therefore, a secondary method for collating the information was chosen. Moreover, all the information was collected from secondary resources such as past articles and research papers [2]. The collection of secondary information aids in collecting reliable information thus a reliable result can be formulated [3]. Furthermore, for analysing the secondary information a qualitative method of analysis was followed. The process of analysing physical education teaching and learning for Chinese students is required to analyse sociological factors. Moreover, physical qualitative methods aid in establishing reliable relations that are routed in society [4]. Therefore, the method of quantitative analysis and the use of a secondary qualitative method is justified for analysing physical education teaching and learning in China.

**3.1. Rules and Routine for PE in China.** At the time of past analysis, it was noted that there are certain rules and related reasons for the rules for PE education in China. For instance, mandatory PE education is one of the major rules that enforce PE education for everyone. In addition, the implication of PE education needs to be safe and diversified for each of the students [5].

From the figure 3.1, the implication process of PE education for Chinese students is described. Moreover, the implication of PE training is described in the process [6]. It is seen that the educational process of PE is dedicated to the development of interest in sports activity.

**3.2. Factors impacting the implementation of rules and routines in PE teaching and learning in China.** To establish a coherent understanding related to the implementation of rules and routines in PE teaching in China understanding the related factors is essential [7]. Moreover, the factors of PE education are subjective and can vary based on situations [8]. Following are some of the essential elements that contribute to the personal nature of PE education.

*Priorities of the education system.* For PE training, the emphasis on academic success and another curriculum can result in less time and funding being allotted for physical education. This may affect how rules and procedures are established and followed in physical education classrooms.

*Chinese Culture and Beliefs.* China's social norms, such as amenability to authority and self-control, can affect the implication of rules for physical education settings [9]. Further, Students' conduct and attitudes

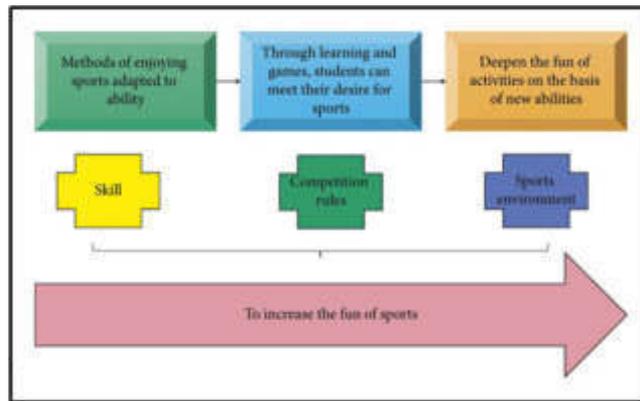


Fig. 3.1: PE education system implication process for Chinese students

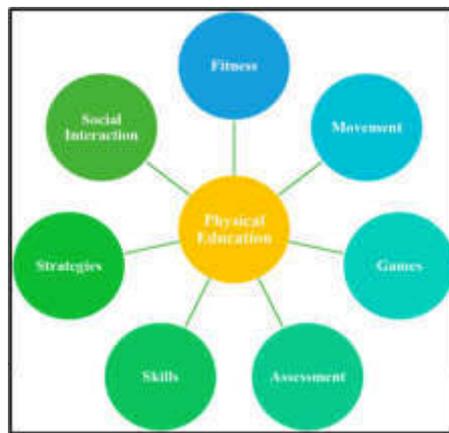


Fig. 3.2: Impact of PE on students

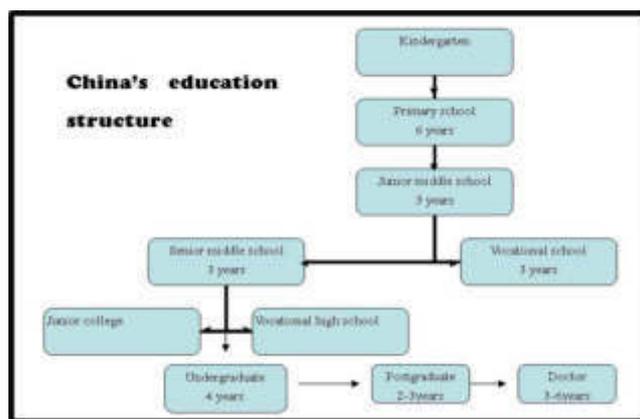


Fig. 3.3: Illustration of the Chinese education system

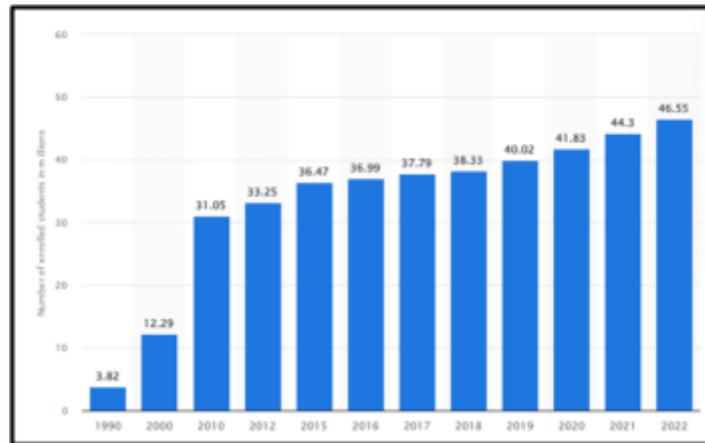


Fig. 3.4: Number of students in tertiary education in China

toward following regulations might be influenced by cultural norms. Thus, it can be contemplated that cultural norms impact the implication of PE teaching in China.

*Availability of resources.* In order to carry out some routines and activities, sports facilities and equipment must be available and adequate. The range of physical education exercises may be limited by a lack of adequate facilities [10].

*Policies made by the government.* Government-imposed educational policies have an impact on how much focus is given to physical education in the curriculum [11]. Rules and procedures can be implemented more successfully with the help of supportive policies.

*Education leadership.* After the government leadership in educational institutions has the authority to implement the rules and regulations for students. Therefore, effective implementation of rules for PE Education can be greatly influenced by the school leadership [12]. Moreover, the dedication to promoting physical education and fostering a supportive atmosphere for rules and routines have a direct relation with educational leaders.

**3.3. Issues related to the rules and regulations for physical education teaching and learning in China.** During the analysis of secondary information related to the rules and regulations related to PE education for Chinese students, some problems were noted. Some of the problems are the following:

*Scarceness of the faculty and resource.* China is one of the majorly populated counties in the world that have a more number of students in schools and colleges. As per the information for 2022 around 59.6% growth in the number of students enrolled in the tertiary education system was recorded [13].

However, from the figure 3.4, it can be seen that in 2021 the number of students was 44.3 billion and it grew to 46.5 billion. On the other hand, a 1.69% in resource allocation was observed in the Chinese education system, thus appropriate allocation of resources is a major issue in maintaining rules and regulations regarding PE education [14].

*Massive class size and limited teachers.* It is difficult for teachers to successfully manage and provide each student with individualized attention during PE courses in Chinese schools due to the prevalence of large class numbers [15]. Such factors can impact the effectiveness of education and the growth of personal skills.

*Gender Inequalities.* In PE classes, there may be a gender imbalance in the activities available, with some activities having a stronger male or female bias. As a result, there may be fewer possibilities for skill development between genders, and gender stereotypes are reinforced [16].

From the figure 3.5, it can be contemplated that there is a growth in the number of students per teacher. It can be noted that the average has gone to 18.5 students per teacher. Considering the large population of China even a growth of 1 on average can be challenging for teachers [17]. Thus, it can be comprehended that there is an issue in human resources that impacts the implications of rules and regulations in physical education teaching and training.

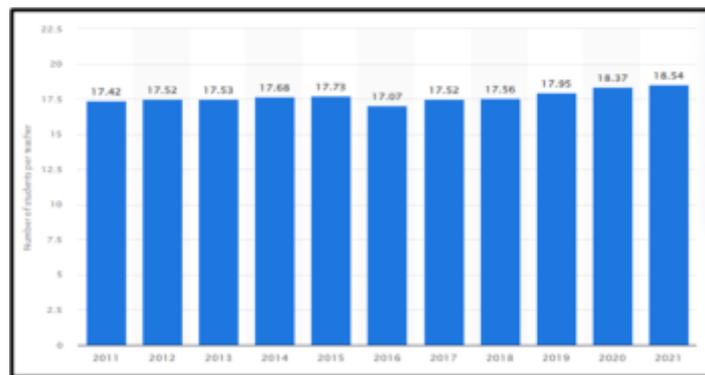


Fig. 3.5: Number of students per teacher in China

**4. Possible solution for countering Issues in the implementation of rules and routines in PE teaching in China.** At the time of analysing past literature for comprehending the implementation of rules and regulations, it was noted that there are certain impacts on the students. Thus, all of the issues need to be countered effectively in order to achieve optimal results for the students. In order to counter the issues following aspects can be followed:

*Allocating appropriate resources for PE.* During the systemic analysis of the issues related to PE education, it was noted that resource allocation is a major problem. Activist related to PE aids to meet the physical as well as mental needs of the students [18]. Thus, allocating resources can benefit the process of implementing rules and routines for students.

*Setting two-tier enforcement.* It was noted that government and school administration are primarily responsible for the implication of PE for students [19]. Therefore, it can be contemplated that with a tow tire administration for the PE teaching basic issues of the system can be resolved. Moreover, resource allocation can be done effectively.

*Diversifying the PE programs.* For the effective implication of the PE programs for student's diversification can be beneficial. Moreover, each student are different in nature and each class has different requirements regarding PE education [20]. Therefore, with the implication of a diversified system, students can be benefited and the shortage of teachers can be countered.

**5. Results and Discussion.** At the time of analysing past literature related to PE education in China, it was noted that there are some rout benefits for the students [21]. Moreover, implementing certain rules and making them an integral part of the education system can aid students in the long term [22]. Additionally, it was noted that there are some central issues such as resource selection and availability of teachers. The issue of location resources in an effective manner hinders the implementation of mandatory rules for the students [23].

The figure 5.1 illustrates the budget allocation per student in China. It can be seen that there are 40% of students receive 1000 to 3000 Yuan per student. Therefore, it can be said that there with such an increase in the budget a better PE education.

On the other hand, it was noted that there are issues related to teachers' availability. It was contemplated that there in order to appropriately understand student needs an increasing number of students was noted [24]. Considering the benefits of PE education for students it can be said that increasing human resources can be beneficial for the quality implication of rules and regulations human resource is an essential factor. Furthermore, it was noted that rules and regulations are essential for the holistic development of the students [25]. Additionally, the rules and routines are set by the education ministry of the government and school authorities [26]. Therefore, it was found that establishing a two-tier regulation can benefit the process of implementing rules and ruins for the students. In such a manner, individual needs of the students can be catted in a systematic manner [27].

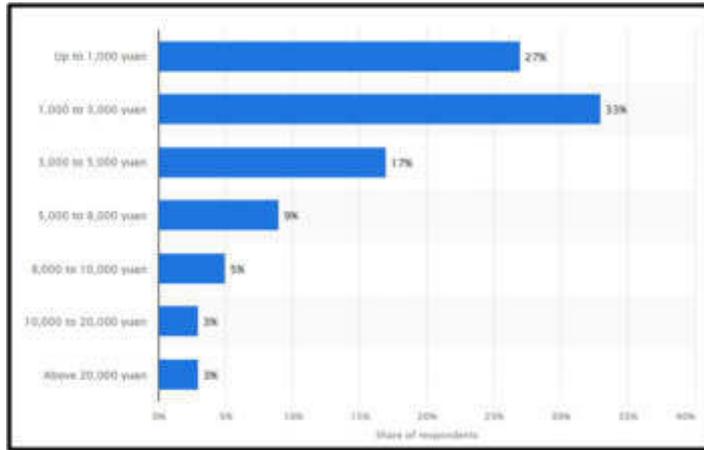


Fig. 5.1: Budget allocation per student in China

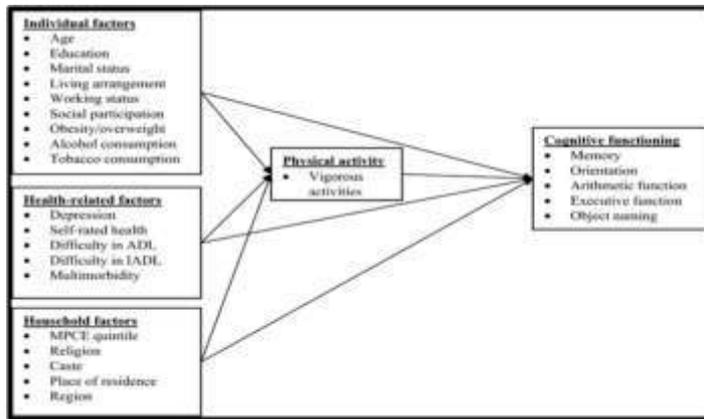


Fig. 5.2: Relation between PE and cognitive ability of students

**6. Conclusion.** Thus, secondary qualitative research was conducted in order to understand the implication of rules and routines for physical education teaching in China. It was noted that PE aid in the overall development of the students. Furthermore, there are physical and mental benefits to PE. On the other hand, it was noticed that there are some issues that hinder the implication of rules for PE. However, it was noted that the implication of resources and diversification in the activities such issues can be countered. By examining the multifaceted aspects of physical education in China, this research contributes to the broader discourse surrounding the role of PE in shaping well-rounded individuals. It sheds light on how China’s educational system incorporates physical education to promote both physical fitness and character development, nurturing a generation of students equipped not only with academic knowledge but also with the tools for a healthy, balanced life.

REFERENCES

[1] D. BAYANOV, L. NOVITSKAYA, S. PANINA, Z. PAZNIKOVA, E. MARTYENKO, K. ILKEVICH, V. KARPENKO, AND R. ALLALYEV, *Digital technology: Risks or benefits in student training?*, Journal of Environmental Treatment Techniques, 7 (2019), pp. 659–663.

[2] F. CAO, M. LEI, S. LIN, AND M. XIANG, *Application of artificial intelligence-based big data ai technology in physical education*

- reform.*, Mobile Information Systems, (2022).
- [3] N. CHEN, C. G. KIM, K. J. J. LEE, J. KIM, N. CHEN, C. G. KIM, K. J. J. LEE, AND J. KIM, *The moderating effect of urbanization on the association between socioeconomic status and physical activity in chinese adults: A cross-sectional study*, Exercise Science, 30 (2021), pp. 288–294.
  - [4] S. M. CHOI, K. W. R. SUM, T. L. WALLHEAD, F. L. E. LEUNG, S. C. A. HA, AND H. P. C. SIT, *Operationalizing physical literacy through sport education in a university physical education program*, Physical Education and Sport Pedagogy, 27 (2022), pp. 591–607.
  - [5] B. DAPENG ET AL., *Research on “smart classroom” teaching mode of public physical education in colleges and universities based on “internet+”*, The Frontiers of Society, Science and Technology, 2 (2020).
  - [6] I. DEMCHENKO, B. MAKSYMCHUK, V. BILAN, I. MAKSYMCHUK, AND I. KALYNOVSKA, *Training future physical education teachers for professional activities under the conditions of inclusive education*, BRAIN. Broad Research in Artificial Intelligence and Neuroscience, 12 (2021), pp. 191–213.
  - [7] X. ERYONG AND J. LI, *What is the ultimate education task in china? exploring “strengthen moral education for cultivating people” (“li de shu ren”)*, Educational Philosophy and Theory, 53 (2021), pp. 128–139.
  - [8] J. A. EZEUGWU, *Computer-based training and web-based learning strategies needed for effective teaching and learning of physical education in secondary schools in nsukka local government*, Journal of Science & Computer Education, 3 (2015).
  - [9] Y. GU, *Deep integration of physical education and multimedia technology using internet of things technology*, Wireless Communications and Mobile Computing, 2022 (2022).
  - [10] Y. GUO, D. WANG, AND X. CHENG, *Integration of information technology and traditional taekwondo curriculum*, International Journal of u-and e-Service, Science and Technology, 7 (2014), pp. 163–172.
  - [11] P. HASTIE, A. HU, H. LIU, AND S. ZHOU, *Incorporating sport education within a physical education sports club in china*, Curriculum Studies in Health and Physical Education, 11 (2020), pp. 129–144.
  - [12] L. HE, Y. CAO, AND J. MAO, *Exploring college students’ fitness and health management based on internet of things technology*, Journal of High Speed Networks, 28 (2022), pp. 65–73.
  - [13] F. J. HINOJO LUCENA, J. LOPEZ BELMONTE, A. FUENTES CABRERA, J. M. TRUJILLO TORRES, AND S. POZO SANCHEZ, *Academic effects of the use of flipped learning in physical education*, International journal of environmental research and public health, 17 (2020), p. 276.
  - [14] T.-C. KANG, C.-H. WEN, S.-W. GUO, W.-Y. CHANG, AND C.-L. CHANG, *The implementation of an iot-based exercise improvement system*, The Journal of Supercomputing, 76 (2020), pp. 6361–6375.
  - [15] M. H. LI, R. K. W. SUM, C. H. P. SIT, S. H. S. WONG, AND A. S. C. HA, *Associations between perceived and actual physical literacy level in chinese primary school children*, BMC Public Health, 20 (2020), pp. 1–9.
  - [16] Y. LIANG, H. GUO, AND H. YI, *Use iot in physical education and sport in china schools*, Wireless Communications and Mobile Computing, 2022 (2022).
  - [17] W. LUO, *Using big data technology to study the countermeasures to improve the teaching ability of pe teachers in applied universities in china*, in Journal of Physics: Conference Series, vol. 1744, IOP Publishing, 2021, p. 042007.
  - [18] C. MANATHUNGA, M. SINGH, J. QI, AND T. BUNDA, *Using chinese and first nations philosophies about time and history to reimagine transcultural doctoral education*, Discourse: Studies in the cultural politics of education, 44 (2023), pp. 121–132.
  - [19] R. MARTIN-SMITH, A. COX, D. S. BUCHAN, J. S. BAKER, F. GRACE, AND N. SCULTHORPE, *High intensity interval training (hiit) improves cardiorespiratory fitness (crf) in healthy, overweight and obese adolescents: a systematic review and meta-analysis of controlled studies*, International journal of environmental research and public health, 17 (2020), p. 2955.
  - [20] G. M. MU, *Chinese education and pierre bourdieu: Power of reproduction and potential for change*, 2020.
  - [21] O. NAFTALI, *Celebrating violence? children, youth, and war education in maoist china (1949–1976)*, The Journal of the History of Childhood and Youth, 14 (2021), pp. 254–273.
  - [22] W. O’BRIEN, M. ADAMAKIS, N. O’BRIEN, M. ONOFRE, J. MARTINS, A. DANIA, K. MAKOPOULOU, F. HEROLD, K. NG, AND J. COSTA, *Implications for european physical education teacher education during the covid-19 pandemic: a cross-institutional swot analysis*, European Journal of Teacher Education, 43 (2020), pp. 503–522.
  - [23] B. PANG, *Engaging bourdieu’s habitus with chinese understandings of embodiment: Knowledge flows in health and physical education in higher education in hong kong*, Educational Philosophy and Theory, 52 (2020), pp. 1256–1265.
  - [24] ———, *The postmonolingual turn: rethinking embodiment with new confucianism in bodily education and research*, Sport, Education and Society, 27 (2022), pp. 893–905.
  - [25] M. SABBAGH, C. SADOWSKY, B. TOUSI, M. E. AGRONIN, G. ALVA, C. ARMON, C. BERNICK, A. P. KEEGAN, S. KARANTZOULIS, E. BAROR, ET AL., *Effects of a combined transcranial magnetic stimulation (tms) and cognitive training intervention in patients with alzheimer’s disease*, Alzheimer’s & Dementia, (2019).
  - [26] B. TU, W. FU, AND Y. TAI, *Implementation of rules and routines in physical education in china through iot systems*, Wireless Communications and Mobile Computing, 2022 (2022).
  - [27] H. YU AND Y. MI, *Application model for innovative sports practice teaching in colleges using internet of things and artificial intelligence*, Electronics, 12 (2023), p. 874.

*Edited by:* Sathishkumar V E

*Special issue on:* Scalability and Sustainability in Distributed Sensor Networks

*Received:* Aug 23, 2023

*Accepted:* Oct 29, 2023



## RESEARCH ON THE DESIGN OF INTEGRATED ENERGY MANAGEMENT AND OPTIMIZATION CONTROL SYSTEMS FOR NOVEL POWER SYSTEMS

ZHIQIAN YANG, XIANYOU WU, QIUHUA CHEN, AISIKAER, LIANGNIAN LV, AND LEI WU\*

**Abstract.** This research paper is determining the impact of integrated energy management and optimization control for Novel power systems. Research objectives have been made based on integrating power system for reducing power consumption based on optimization parameters. Optimal control is mainly dealing with control law by finding out a given system and delivering certain criteria for achieving goals. There are mainly three parts for problem optimization purposes those are included decision, constraints and objective functions. Four different objectives has been made those are mainly discussed about the importance of the power systems regarding integrated management system. There are various factors that are affecting optimizations including the role of the multigrain recursion, global convergence, properties of the optimization model and local convergence. Along with this, the optimization control system has been played a crucial role for the power systems development purposes. Research has been based on the effective design of energy management for power systems. The primary application for the optimization techniques is based on the storage system, electromagnetic-based design and mapping design for the microwave structure. The rate of unsustainable energy management is increasing and sustainable energy management system is decreasing. The conclusion has been based on the significance of the optimization control for the power systems.

**Key words:** Global convergence, Novel power systems, integrated energy management, optimization control, Problem optimization, local convergence, Energy management

**1. Introduction.** An energy management system or EMS is a system that is based on a computer-aided tool and it is mainly used by the operators of the electric utility grids for controlling, monitoring and performance for generating and determining the transmission system. Management optimizations are delivering efficient uses of energy and delivering the best performances for production purposes. There are various effective strategies for energy management purposes those are includes tracking progress, collection of bill data, and identifying sources and the rate of energy consumption.

The growing demand for efficient and sustainable energy management systems has become a pressing concern in the contemporary world. To address this concern, this research delves into the profound impact of integrated energy management and optimization control within novel power systems. By investigating this critical intersection of energy management and control, we aim to shed light on how such integration can lead to reduced power consumption and, consequently, a more sustainable energy landscape.

Optimal control, as a central theme of this research, focuses on the development and implementation of control laws that are meticulously tailored to a given system. These control laws play an instrumental role in guiding the system toward achieving specific criteria and goals. The optimization process itself is a multifaceted endeavor, encompassing decision-making, constraints, and objective functions, which jointly shape the trajectory of power systems.

### 2. Objectives.

1. To determine the importance of an energy management system for the power system
2. To evaluate the importance of energy optimization for the Novel power systems
3. To examine the key components of integrated energy management for the Novel power system
4. To analyze the optimization technique in the Novel power system.

Our research revolves around four distinct objectives, each designed to highlight the pivotal importance of power systems in the context of an integrated management framework. These objectives delve into a range of critical factors that influence optimization outcomes, including the multifaceted aspects of multigrain recursion,

---

\*Goldwind Science & Technology Co., Ltd, Beijing, China, 100176 ([aisikearresearch1@outlook.com](mailto:aisikearresearch1@outlook.com))

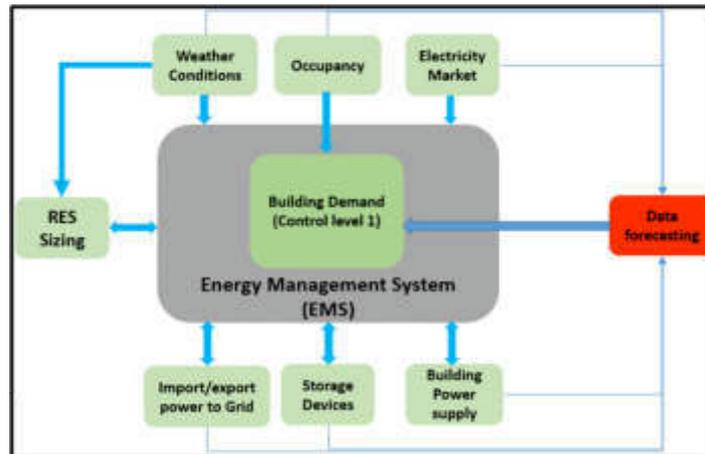


Fig. 3.1: Overview of EMS

the global and local convergence properties of the optimization model, and the role of power systems in this grand scheme.

**3. Methodology.** The analysis procedure is based on a wider spectrum of knowledge-based analytic approaches determining purposes. Secondary data analysis has been done in this research paper and this is determining renewable energy and numerical information for the research paper. Secondary data analysis has a higher scope on developing renewable energy for determining Novel power systems. Qualitative data has been generated based on valuable outcomes for determining the reliability and efficiency of energy resources [1].

**3.1. Importance of an energy management system for the power system.** The major purpose of BEMS is to determine the optimization of energy and energy management that is mainly based on information analysis. The energy management system is decreasing the overall costs of the operations and it is developing the rate of productivity [2].

**Functions of EMS:** EMS is mainly referring to the collection of automated tools that are mainly used for monitoring, controlling and optimizing the overall performances, sub transmissions systems. The significant EMS includes energy targets, measurements, energy policy, roles and responsibilities and energy efficiency development plan. For cost-saving purposes and increasing reputations and competitive edge, EMS is playing a significant role [3]. It is delivering various stages for decreasing electrical energy consumption costs and people can save around 30% on total energy costs with the help of the implementations of the EMS. The main purpose of the EMS is to manage and achieve optimum energy utilization, and procurements with the help of the organizations. It is helping to decrease energy costs and environmental effects [4].

The figure 3.1 is representing the impact of the EMS, which is based on various aspects those are included occupancy, storage device, building power supply and RES sizing for data forecasting purposes.

**Reduce operational costs and productivity:** The major benefit of EMS's ability is to decrease the overall electricity costs by optimizing energy and monitoring. With the help of the data collection, administrations are predicting the usage of energy usage and effective budget handling purposes [5]. Lighting and temperature regulation is playing a significant role in developing productivity and it is helping to maintain indoor temperature along with minimal lights. EMS is a system for optimizing, controlling and monitoring transmission and energy usage and it is determining the ability of the heating, ventilation and lighting [6]. It is helping to decrease the materials and time waste and deleting the employee's perks and determining efficiency equipment purposes.

**Build a positive brand image:** Organizations are trying to decrease the carbon footprint rate by decreasing the energy consumption rate. It is helping in better relationship management purposes among potential investors and partners [7]. It is creating value for the functional aspects and develops emotional

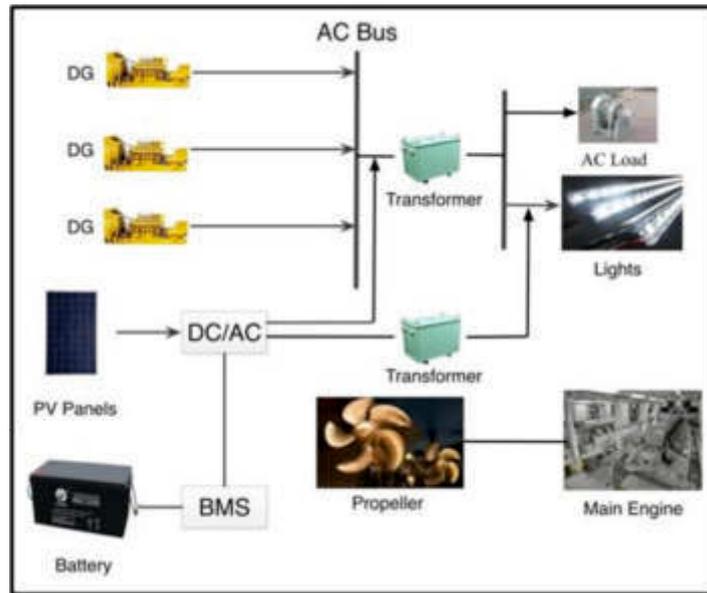


Fig. 3.2: Novel energy management strategy

conditions for the development of loyalty and competition. There are various ways for creating positive and effective brand images those are includes delivering customer services, creating valuable content, search engine results analysis and understanding the target market.

**Increase ROI and Increase property value:** Installing EMS is needed for cost-effective solutions and people are optimizing this system for conserving energy and delivering ideal solutions for global warming [8]. Power transmission is regulated by the EMS and it is expanding the lifespan by increasing the ROI rate and appliances. Satisfying customers and delivering efforts for the development of property management purposes, it is playing a significant role. ROI is helping to develop revenue arte and sales growth for the organizations and developing digital marketing aspects and delivering contributions for revenue development purposes [9]. Capitalizing the market upswings, and reinvesting gains, is playing a significant role. ROI is playing a significant role for developing the profit rate and developing invested property.

**3.2. Significance of energy optimization for the Novel power systems.** Optimizing energy systems is developing efficiency and cost-effectiveness for saving money and decreasing greenhouse gas emissions. It is creating jobs; meeting the demands and saving money that are the major importance for energy systems development purposes [10]. Effective energy optimization can lead to significant cost savings. Power systems often account for a substantial portion of operational expenses, and optimizing energy consumption can reduce these costs. It also helps in prolonging the life of equipment and reducing maintenance and replacement expenses.

**Purpose and benefits of the energy system:** For the industrial process energy optimization is playing a significant role and developing energy efficiency rate. The major key components for reaching the goals are to develop process integration techniques and focus on the heat exchanges that are mainly used in the industry. The major purpose of energy control programs is to manage safety in the workplace and prevent start-up and motion [11].

**Optimization techniques in power systems:** There are various traditional methods for the power system those are included QP, NLP and NFP. There are mainly four stages for the optimization those are included analysis, research, and implementation purposes. Global optimization methods and local optimization methods are the major methods for optimization purposes [12].

The figure 3.2 is representing the strategy for the novel energy system that is based on BMS and DC or AC

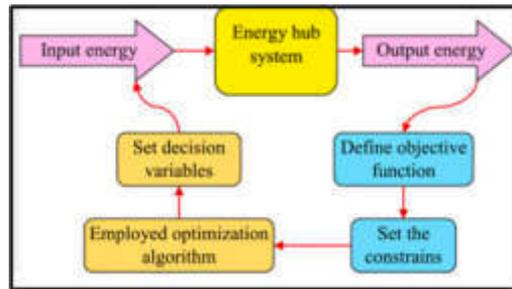


Fig. 3.3: Economic programming for an energy hub in the power system

system. The key features of the optimization are based on decision variables, objective functions and business constraints [13]. Along with this, there are mainly two types of optimization those are included discrete and continuous optimization. The positive impact of the optimizations includes accurate information, higher quality of results, developing efficiency and greater adaptability [14].

The figure 3.3 is representing the overall process for the input and output energy for the energy hub system. It is based on four different stages those are included setting constraints, proper decisions, defining objective functions and employing optimization algorithms [15].

**3.3. Key components of integrated energy management for the Novel power system.** Key components for energy management are based on communication network, measurement devices and software applications. Those systems are helping to manage information efficiently and based on various aspects for an effective EMP [16]. There are mainly six stages for the EMS which are mainly based on getting proper commitment, developing an action plan, understanding the issues, planning and reporting. Environmental sustainability, social, human and economic sustainability are the major key component for sustainability development purposes [17]. Four major key features for the energy-efficient design mainly consist of natural light, solar gains and insulation [18].

A robust communication network serves as the backbone of integrated energy management. It facilitates real-time data exchange between various components of the power system. This network enables seamless monitoring, control, and coordination of energy resources and loads. Moreover, it ensures that the system can respond promptly to dynamic changes in demand and supply. Power systems are part of a broader energy grid. Energy optimization contributes to grid stability by preventing overloads and power imbalances. This is especially important as the grid integrates more renewable energy sources.

**4. Challenges faced by the Novel power supply.** Computations are needed non-zero time for determining the uncertainty and truncation is negatively affecting stability. The primary challenges for the optimizations are include workloads, lack of proper resources, visibility and remote work [19]. Along with this, there are various challenges for the power system those are including threats from cyber attacks, grid modernizations, frequency power outages and electricity transmission losses [20]. The major problems of the power system include voltage swell, dips, voltage unbalance and harmonic. The shortage of fuel, differential tariff structure, delays in tariff revisions and higher AT&C losses. The future challenges for the power system are mainly based on cyber challenges, environmental emissions, electricity load growth and environmental emissions and climate change [21]. Power outages are the common cause of power system failure purposes and it has occurred due to human error, equipment malfunctions, storms and surges in the system. Inefficient coal supplies in thermal power plants are creating a power crisis in the power points [22]. There are four reasons for the poor power quality those are includes load imbalance, harmonic pollution, low power factors and voltage variations.

**5. Results.** The positive impact of this kind of algorithm is to determine the control of DGs and this algorithm is solving single point chain issues. With the help of the virtual leader, this algorithm is directly setting a proper design and achieving frequency and output for achieving leader. It is helping to design the secondary control data and it is following basic theory.

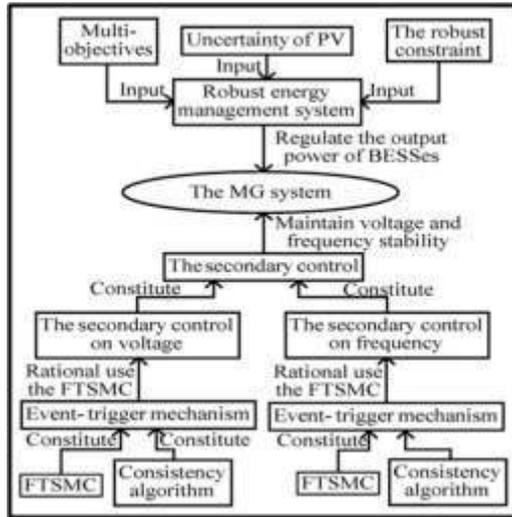


Fig. 5.1: Integrated design for energy and control management

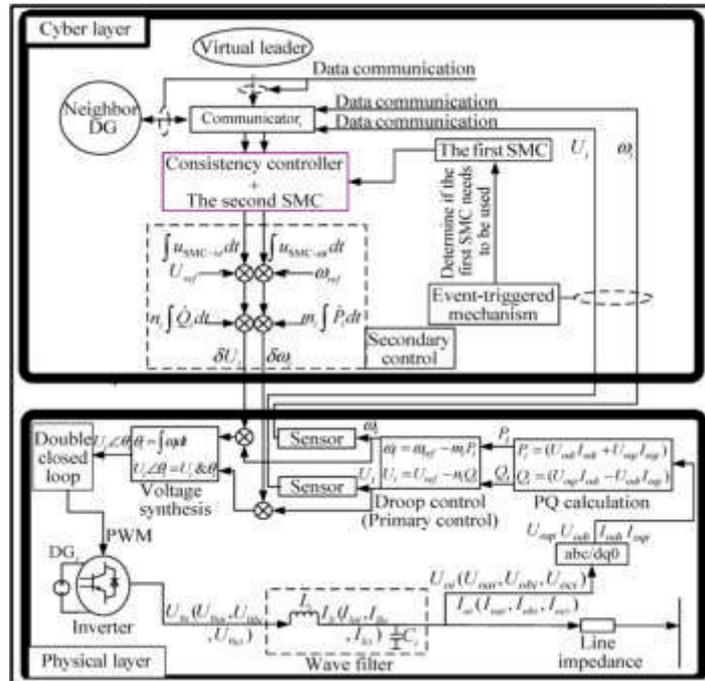


Fig. 5.2: Structure for control structure

The figure 5.1 is representing an integrated design that is mainly based on two different strategies those are included a robust optimal management strategy ad secondary control strategy. This is based on consistency controller, mechanism and FTSMC and for the energy management system; it is based on robust and multi objectives constraints [23].

The figure 5.2 is showing the designed control structure that is based on the physical and cyber layer. The voltage and cyber layer is the most common layer that is affecting the main grid. The frequency and

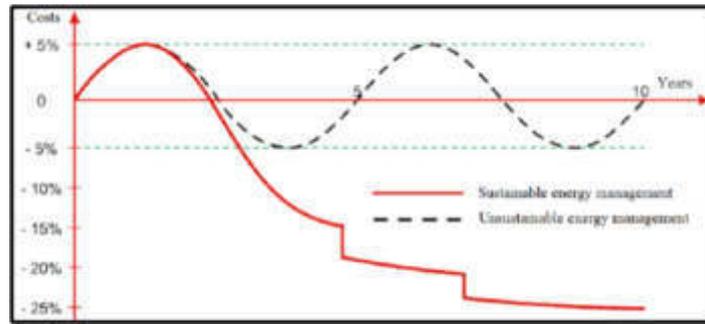


Fig. 5.3: Graph for sustainable energy management system

output voltage have been determined by the droop control and the droop control is designed by the generating operations [23]. In this research study, the designed secondary data control methods are needed for the DG's controlling purposes.

The novel power systems are delivering power efficiency, greater performance and better quality results for the power systems. It is mainly specialized in designing and manufacturing electrical power systems and all the products are based on higher and advanced technology. It is mainly a microprocessor and analog planning-based power system and all the products are measured by Mil-std1275 and ISO [24].

The figure 5.3 is showing the effectiveness and the graph chart for the sustainable energy management system. The functions of the EMS are includes security assessment, network model building, dispatch and automatic generation control. Integrated energy management is helping to schedule, forecast, commercial statements and accounting purposes. Those are helping in proper interaction along with renewable energy and delivering competition [25]. EMS is decreasing the rate of energy consumption, developing industrial productivity, and improving global enterprise. IES is developing cost effectiveness and modern and advanced temperature is heating and developing efficiency.

The energy transition is influencing the development of renewable energy sources and the overall growth of the energy production is creating complexity regarding grid managers and regulating electricity in the supply chain. Result analysis has been showing that smart meter data is delivering proposed machine learning and block chain-based energy trading purposes [26]. The power management method for the multi-source systems mainly leads to the application of the charging stations and delivering optimal operation planning for cost-considering purposes. Integrated energy system is developing the quality of the grid and it is maintaining the balancing power and capacity. There are various ways for controlling the power system those are included regulations, using shunt capacitors, changing transformations and controlling voltage [27].

For developing stability, and reliability for the power system, energy storage systems are playing a significant role in the power system. Cost saving is the major benefit of energy management and it is decreasing climate change and developing reputations. The crucial components of the EMS are includes energy policy, roles and responsibilities, measurements, monitoring, energy targets and objectives [28]. The power system, it is playing a significant role in operating electric utilities for managing, optimizing, controlling and delivering proper performances for the transmission system.

**6. Conclusion.** The research findings presented in this study have far-reaching implications for the development of power systems and energy management. The study primarily focuses on the design and implementation of integrated energy maintenance and optimization control systems for Novel power systems. One of the standout results of this research is the development of an algorithm that effectively controls Distributed Generators (DGs). This algorithm not only addresses single point chain issues but also utilizes a virtual leader, which plays a pivotal role in establishing a robust design and achieving precise control of frequency and power output. Such advancements in control algorithms are essential for improving the reliability and efficiency of power generation from DGs, which are increasingly integrated into modern power systems.

Furthermore, the research introduces an integrated design that combines a robust optimal management

strategy with a secondary control strategy. These strategies are essential for maintaining grid stability and efficient power management. The incorporation of elements like consistency controllers, mechanisms, Fault-Tolerant Sliding Mode Control (FTSMC), and multi-objective constraints demonstrates the complexity and depth of this integrated design. This research also underscores the necessity of well-defined secondary data control methods, especially in the context of DGs, where precise control is crucial for grid reliability. This research offers a comprehensive understanding of the integrated energy maintenance and optimization control systems for Novel power systems. The findings demonstrate the critical role of advanced algorithms, integrated designs, and sustainable energy management systems in the modern power landscape. As the energy sector continues to evolve, these findings are pivotal in guiding the development of more efficient, reliable, and sustainable power systems.

**Fundings.** This work is funded by the Science and Technology Department of Xinjiang Uygur Autonomous Region under grant No. 2022A01004-2.

#### REFERENCES

- [1] S. E. AHMADI AND N. REZAEI, *A new isolated renewable based multi microgrid optimal energy management system considering uncertainty and demand response*, International Journal of Electrical Power & Energy Systems, 118 (2020), p. 105760.
- [2] T. S. BABU, K. R. VASUDEVAN, V. K. RAMACHANDARAMURTHY, S. B. SANI, S. CHEMUD, AND R. M. LAJIM, *A comprehensive review of hybrid energy storage systems: Converter topologies, control strategies and future prospects*, IEEE Access, 8 (2020), pp. 148702–148721.
- [3] A. BARZKAR AND M. GHASSEMI, *Electric power systems in more and all electric aircraft: A review*, Ieee Access, 8 (2020), pp. 169314–169332.
- [4] D. CAO, W. HU, J. ZHAO, G. ZHANG, B. ZHANG, Z. LIU, Z. CHEN, AND F. BLAABJERG, *Reinforcement learning and its applications in modern power and energy systems: A review*, Journal of modern power systems and clean energy, 8 (2020), pp. 1029–1042.
- [5] V. DUDJAK, D. NEVES, T. ALSKAIF, S. KHADEM, A. PENA-BELLO, P. SAGGESE, B. BOWLER, M. ANDONI, M. BERTOLINI, Y. ZHOU, ET AL., *Impact of local energy markets integration in power systems layer: A comprehensive review*, Applied Energy, 301 (2021), p. 117434.
- [6] A. FERNÁNDEZ-GUILLAMÓN, E. GÓMEZ-LÁZARO, E. MULJADI, AND Á. MOLINA-GARCÍA, *Power systems with high renewable energy sources: A review of inertia and frequency control strategies over time*, Renewable and Sustainable Energy Reviews, 115 (2019), p. 109369.
- [7] J. GUERRERO, D. GEBBRAN, S. MHANNA, A. C. CHAPMAN, AND G. VERBIČ, *Towards a transactive energy system for integration of distributed energy resources: Home energy management, distributed optimal power flow, and peer-to-peer energy trading*, Renewable and Sustainable Energy Reviews, 132 (2020), p. 110000.
- [8] Q. HUANG, R. HUANG, W. HAO, J. TAN, R. FAN, AND Z. HUANG, *Adaptive power system emergency control using deep reinforcement learning*, IEEE Transactions on Smart Grid, 11 (2019), pp. 1171–1182.
- [9] X. LI AND S. WANG, *Energy management and operational control methods for grid battery energy storage systems*, CSEE Journal of Power and Energy Systems, 7 (2019), pp. 1026–1040.
- [10] Y. LI, M. HAN, Z. YANG, AND G. LI, *Coordinating flexible demand response and renewable uncertainties for scheduling of community integrated energy systems with an electric vehicle charging station: A bi-level approach*, IEEE Transactions on Sustainable Energy, 12 (2021), pp. 2321–2331.
- [11] Y. LI, C. WANG, G. LI, J. WANG, D. ZHAO, AND C. CHEN, *Improving operational flexibility of integrated energy system with uncertain renewable generations considering thermal inertia of buildings*, Energy Conversion and Management, 207 (2020), p. 112526.
- [12] T. LIU, X. TANG, H. WANG, H. YU, AND X. HU, *Adaptive hierarchical energy management design for a plug-in hybrid electric vehicle*, IEEE Transactions on Vehicular Technology, 68 (2019), pp. 11513–11522.
- [13] Z. LIU, A. MOHAMMADZADEH, H. TURABIEH, M. MAFARJA, S. S. BAND, AND A. MOSAVI, *A new online learned interval type-3 fuzzy control system for solar energy management systems*, IEEE Access, 9 (2021), pp. 10498–10508.
- [14] M. MAHZARNIA, M. P. MOGHADDAM, P. T. BABOLI, AND P. SIANO, *A review of the measures to enhance power systems resilience*, IEEE Systems Journal, 14 (2020), pp. 4059–4070.
- [15] M. MIR, M. DAYYANI, T. SUTIKNO, M. MOHAMMADI ZANJIREH, AND N. RAZMJOOY, *Employing a gaussian particle swarm optimization method for tuning multi input multi output-fuzzy system as an integrated controller of a micro-grid with stability analysis*, Computational Intelligence, 36 (2020), pp. 225–258.
- [16] S. M. MOGHADDAS-TAFRESHI, M. JAFARI, S. MOHSENI, AND S. KELLY, *Optimal operation of an energy hub considering the uncertainty associated with the power consumption of plug-in hybrid electric vehicles using information gap decision theory*, International Journal of Electrical Power & Energy Systems, 112 (2019), pp. 92–108.
- [17] C. MOKHTARA, B. NEGROU, A. BOUFERROUK, Y. YAO, N. SETTOU, AND M. RAMADAN, *Integrated supply–demand energy management for optimal design of off-grid hybrid renewable energy systems for residential electrification in arid climates*, Energy Conversion and Management, 221 (2020), p. 113192.

- [18] W. MUSA, V. PONKRATOV, A. KARAEV, N. KUZNETSOV, L. VATUTINA, M. VOLKOVA, O. SHALINA, AND A. MASTEROV, *Multi-cycle production development planning for sustainable power systems to maximize the use of renewable energy sources*, Civil Engineering Journal, 8 (2022), pp. 2628–2639.
- [19] R. MUZZAMMEL, R. ARSHAD, S. MEHMOOD, AND D. KHAN, *Advanced energy management system with the incorporation of novel security features*, International Journal of Electrical and Computer Engineering, 10 (2020), p. 3978.
- [20] G. PAN, W. GU, Y. LU, H. QIU, S. LU, AND S. YAO, *Optimal planning for electricity-hydrogen integrated energy system considering power to hydrogen and heat and seasonal storage*, IEEE Transactions on Sustainable Energy, 11 (2020), pp. 2662–2676.
- [21] N. PRIYADARSHI, V. K. RAMACHANDARAMURTHY, S. PADMANABAN, AND F. AZAM, *An ant colony optimized mppt for standalone hybrid pv-wind power system with single cuk converter*, Energies, 12 (2019), p. 167.
- [22] S. K. RATHOR AND D. SAXENA, *Energy management system for smart grid: An overview and key issues*, International Journal of Energy Research, 44 (2020), pp. 4067–4109.
- [23] N. WANG, X. ZHOU, X. LU, Z. GUAN, L. WU, X. DU, AND M. GUIZANI, *When energy trading meets blockchain in electrical power system: The state of the art*, Applied Sciences, 9 (2019), p. 1561.
- [24] Y. XIANG, H. CAI, C. GU, AND X. SHEN, *Cost-benefit analysis of integrated energy system planning considering demand response*, Energy, 192 (2020), p. 116632.
- [25] K. Y. YAP, C. R. SARIMUTHU, AND J. M.-Y. LIM, *Artificial intelligence based mppt techniques for solar power system: A review*, Journal of Modern Power Systems and Clean Energy, 8 (2020), pp. 1043–1059.
- [26] R. V. YOHANANDHAN, R. M. ELAVARASAN, P. MANOHARAN, AND L. MIHET-POPA, *Cyber-physical power system (cpps): A review on modeling, simulation, and analysis with cyber security applications*, IEEE Access, 8 (2020), pp. 151019–151064.
- [27] Z. ZHANG, D. ZHANG, AND R. C. QIU, *Deep reinforcement learning for power system applications: An overview*, CSEE Journal of Power and Energy Systems, 6 (2019), pp. 213–225.
- [28] B. ZHOU, J. ZOU, C. Y. CHUNG, H. WANG, N. LIU, N. VOROPAI, AND D. XU, *Multi-microgrid energy management systems: Architecture, communication, and scheduling strategies*, Journal of Modern Power Systems and Clean Energy, 9 (2021), pp. 463–476.

*Edited by:* Sathishkumar V E

*Special issue on:* Scalability and Sustainability in Distributed Sensor Networks

*Received:* Aug 23, 2023

*Accepted:* Oct 29, 2023



## STABILITY STUDY OF NEW POWER SYSTEM BASED ON MULTI-INTELLIGENT BODY COLLABORATION

XIANYOU WU, ZHIQIAN YANG, XIN DU, LIANGNIAN LV, AISIKAER\* AND YANCHEN YANG†

**Abstract.** Developing, implementing, and maintaining a multi-intelligent body collaboration system necessitates significant investments in finances, time, and expertise. While multi-intelligent body collaboration has the potential to enhance power system stability significantly, it also comes with challenges related to interoperability, security, system complexity, and resource allocation. Resource allocation and training costs can be substantial. Addressing these challenges is crucial to harnessing the full benefits of this approach and ensuring the reliable and efficient operation of power systems. Effective communication and coordination strategies among intelligent agents are integral to maintaining power system stability. Timely information exchange, load balancing, disturbance management, and the integration of AI contribute to a more resilient and adaptive energy grid. As technology advances, refining these strategies will be essential to meet the growing demands of an ever-evolving power landscape. As technology marches forward, it becomes increasingly evident that the refinement of these strategies is paramount. The dynamism of the power landscape, driven by technological advancements and evolving needs, necessitates an agile and adaptable power system. The fusion of multi-intelligent bodies and modern technology stands as a testament to our collective pursuit of a more reliable, efficient, and sustainable energy future. In this ever-evolving landscape, the innovation and enhancement of these strategies are our compass, guiding us toward a brighter and more efficient future.

**Key words:** Multiintelligence, Sustainability, Power system, Challenges, Communication, Behaviors.

**1. Introduction.** A stability study of a new power system based on multi-intelligent body collaboration involves assessing the robustness and reliability of this innovative approach to energy management. This system employs multiple intelligent entities that work collaboratively to ensure the stability of the power grid. At its core, this power system harnesses the capabilities of artificial intelligence, machine learning, and advanced control algorithms to manage the intricate balance between power generation, distribution, and consumption. The term "multi-intelligent body collaboration" implies the interaction of various intelligent components, such as smart meters, sensors, control centers, and even consumer devices, in a coordinated manner. One key aspect of this study involves evaluating the system's ability to predict and respond to power demand and supply fluctuations.

The modern power system is a dynamic and complex network that forms the backbone of our modern civilization. It supplies electricity to homes, industries, and institutions, ensuring our daily activities run smoothly. As the electricity demand grows, power systems face increasingly complex challenges. These challenges include the integration of renewable energy sources, the need for improved grid resilience, and the optimization of energy distribution. In this context, a multi-intelligent body collaboration system has emerged as a promising solution to enhance the stability and efficiency of power systems.

While the potential benefits of multi-intelligent body collaboration in power systems are substantial, it is essential to recognize that its implementation is not without challenges. This article delves into the intricacies of this approach, focusing on the key aspects that need to be addressed to exploit its advantages fully. We will explore the challenges related to interoperability, security, system complexity, resource allocation, and the substantial costs associated with training and resource allocation. By understanding and mitigating these challenges, the power industry can harness the full potential of a multi-intelligent body.

### 2. Objective.

1. To analyze the potential benefits and challenges of implementing multi-intelligent body collaboration in power system stability enhancement

---

\*Goldwind Science & Technology Co., Ltd, Beijing, 100176, China (email: [aisikaerres1@outlook.com](mailto:aisikaerres1@outlook.com))

†State Grid Jibei Electric Power Co., Ltd, Beijing, 100032, China

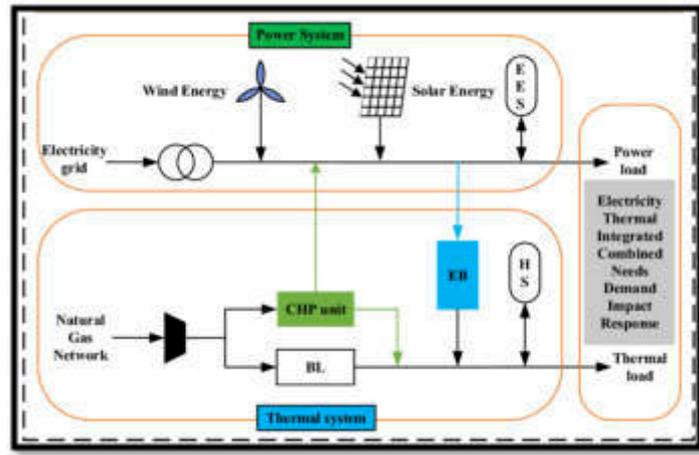


Fig. 3.1: Structure of combined and thermal energy system

2. To develop a comprehensive model of the new power system integrating multi-intelligent bodies
3. To evaluate the impact of different intelligent agent behaviors on power system stability under various operating conditions and disturbances
4. To identify the effectiveness of communication and coordination strategies among intelligent agents in maintaining power system stability

**3. Methodology.** This study mainly uses a secondary qualitative data collection method to gather information related to the work on a new power system based on multi-intelligent body collaboration. Furthermore, the study should assess the economic viability of implementing such a system [2]. The secondary data collection method must be statistical and theoretical, and it can easily give information related to the topic. Additionally, it is essential to evaluate the financial feasibility of putting such a system into practice.

**3.1. Identify the benefits and challenges of implementing multi-intelligent body collaboration in power system stability.** Various potential challenges and benefits are available, and smart grid implication is one of the most influencing challenges. Along with that, there are high investment costs, privacy, operational complexity, and maintaining security; these types of issues might be considered potential challenges [1]. Multi-intelligent body collaboration in power system stability enhancement offers numerous benefits. Firstly, it can leverage the combined capabilities of multiple intelligent agents, such as AI algorithms and control systems, to swiftly detect and respond to stability threats. This can lead to faster fault identification and corrective actions, minimizing disruptions [3]. There are various benefits also available, such as a smart grid that is connected with efficient transmission of electricity.

The above-mentioned Figure 3.1 indicates that this study is mainly focusing on the “wind energy (WT), solar energy (PV), gas turbine (MT), electric heat boiler (EB), gas boiler (BL), energy storage battery (EES), heat storage tank (HS)”, and it is demanding on the side-electric energy heat machine [4]. Additionally, collaboration among various intelligent components allows for enhanced prediction accuracy. By integrating data from diverse sources, like sensors, weather forecasts, and load patterns, the system can make more informed decisions, leading to improved stability and reduced risk of blackouts. However, this approach also presents challenges. Interoperability among different intelligent components is a critical hurdle [5]. Ensuring seamless communication and coordination between AI agents, control systems, and human operators requires sophisticated integration techniques and standardization protocols. Data security and privacy are another concern. Collaborative systems involve sharing sensitive operational data among different intelligent agents. Protecting information from unauthorized access or cyber-attacks demands robust cyber security measures.



Fig. 3.2: Processes for using Multi Intelligences

**3.2. Develop a model of the new power system integrating multi-intelligent bodies.** The model powering system is a combination of analyzing and studying behaviors and performance of the power system. In an era marked by rapid technological advancements and the increasing demand for sustainable energy solutions, the development of a comprehensive model for a new power system that integrates multi-intelligent bodies has garnered significant attention [6]. This study explores the key components and benefits of such a system, highlighting its potential to revolutionize the way we generate, distribute, and consume energy. There are various sources and power systems available, and those are:

*Integration of Renewable Energy Sources.* The new power system model emphasizes the integration of renewable energy sources, such as solar, wind, hydro, and geothermal power. These sources offer a sustainable alternative to traditional fossil fuels, reducing greenhouse gas emissions and mitigating climate change [7].

*Advanced Smart Grid Infrastructure.* Smart grid infrastructure is an electricity network it is used in various technologies, and it helps to monitor and manage those sources that help to fulfill the demand for electricity for users. Smart grids enable real-time communication and coordination between various energy generation and consumption points, optimizing energy distribution and load management [8]. The incorporation of a smart grid infrastructure forms the backbone of the multi-intelligent body-integrated power system.

*Decentralized Energy Generation.* The new model shifts from centralized power generation to decentralized systems. This approach empowers individual households, businesses, and communities to generate their energy using renewable sources, contributing to energy self-sufficiency and grid resilience [9].

Developing a comprehensive model for a new power system that integrates multi-intelligent bodies holds immense promise for transforming our energy landscape. This theory was developed by Howard Gardner in the year 1983 to leave the concept of traditional intelligence and introduce the world to multiple intelligence processes [12]. This model leverages the power of renewable energy sources, advanced technology, data analytics, and distributed AI to create a more sustainable, efficient, and resilient energy infrastructure.

**4. Evaluate the impact of different intelligent agent behaviors on power system stability.** Various impacts are available from different intelligence and power system stability; along with that, Voltage stability is the main contrast that impacts electric system operation. The impact of different intelligent agent behaviors on power system stability is a complex topic that requires careful analysis [10]. Intelligent agents can play a significant role in enhancing the stability of power systems under various operating conditions and disturbances. These agents can include advanced control algorithms, machine learning models, and optimization

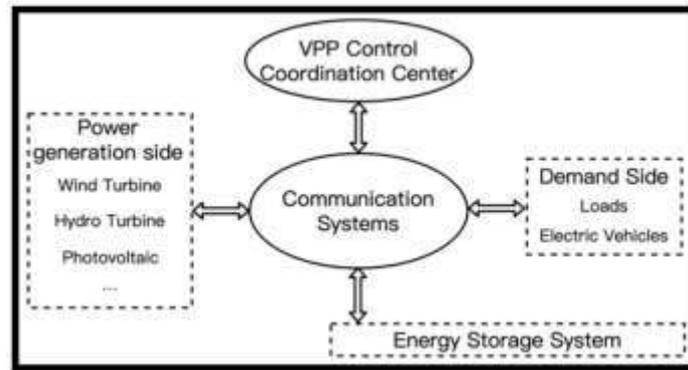


Fig. 4.1: Communication system in renewable energy sources

techniques. Voltage fluctuations might cause voltage stability, which can happen in equipment failures and power outages [11]. One central area where intelligent agents impact is load forecasting and demand response. By accurately predicting future load patterns and adjusting demand accordingly, agents can help maintain a balanced supply-demand relationship, reducing the likelihood of instability caused by sudden load changes.

**4.1. Analyze the effectiveness of communication and coordination strategies in maintaining power system stability.** There is a strong connection available for making good communication, and it plays a significant role in “conveying information for ensuring a stable supply of electricity”. Effective communication and coordination strategies among intelligent agents play a pivotal role in maintaining power system stability [12]. In today’s complex and interconnected energy landscape, where power generation, distribution, and consumption are constantly evolving, the need for seamless information exchange and collaboration cannot be overstated. Firstly, clear communication is crucial for timely decision-making and response. Secondly, effective coordination among agents ensures a balanced load distribution [13]. Furthermore, coordination enhances resilience against disturbances.

Incorporating machine learning and AI techniques into communication and coordination strategies further enhances their effectiveness. These technologies enable agents to predict potential instability based on historical data and emerging patterns. By proactively addressing these concerns, the power system can avoid instability triggers before they occur. However, challenges do exist. Ensuring interoperability among diverse intelligent agents, which may be developed by different manufacturers or for different purposes, requires standardized communication protocols [14]. Cyber security measures are also paramount to protect against malicious attacks that could exploit communication vulnerabilities.

**5. Results.** In the above-mentioned figure 4.1, it is showing the most significant issue in the power grid design and operation are frequency stability and control. There are various processes have been introduced to incorporate this process [15]. When it comes to power reserves, each specification is unique and requires a certain amount of reserve to handle any deviations in power. It’s important to keep these reserves available to ensure smooth operations and prevent any disruptions. When it comes to balancing supply and demand in the energy market, the key lies in controlling the output of dispatch able generating units. This is where most of the work is done, ensuring that energy is being produced and distributed in the most efficient and effective way possible [16].

It’s a complex process, but one that’s absolutely essential to keeping the lights on and the world running smoothly. Moreover, the collaborative nature of these intelligent bodies ensured efficient energy distribution and load balancing. By continuously exchanging data and making informed decisions, the system optimizes energy flow, minimizes wastage, and enables the integration of renewable energy sources at a larger scale [20]. This not only enhances environmental sustainability but also mitigates the challenges posed by fluctuating power generation.

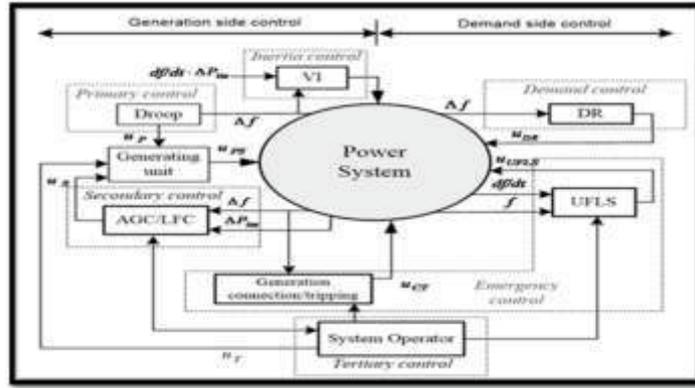


Fig. 5.1: Frequency stability and control smart grid in power system

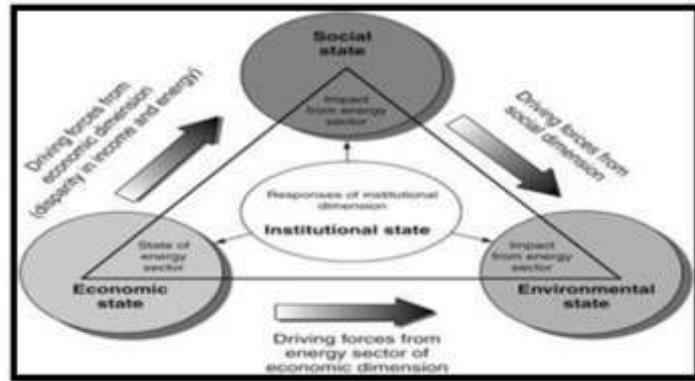


Fig. 5.2: Sustainable energy transaction

The result of a stability study conducted on a novel power system based on multi-intelligent body collaboration yields insights into its potential for revolutionizing energy networks, as mentioned above. This innovative approach leverages advanced technologies like artificial intelligence, machine learning, and interconnected devices to enhance system stability and reliability [17]. The maximum number of people are facing sustainable energy resource issues due to greenhouse gas emissions and global climate change. The stability study shows that integrating multiple intelligent entities, such as smart grids, IoT devices, and predictive analytics, contributes to real-time monitoring, fault detection, and adaptive control. These issues mainly happened because “fossil fuels led by coal, natural gas, and oil contributed 61.3% of global electricity generation in the year 2020”. This synergy reduces the risk of cascading failures and blackout events, improving the overall resilience of the power system [18]. Energy plays a crucial role in all nations’ progress and sustainable development. It is widely recognized as one of the most important factors determining the pace at which a country grows and develops.

Without adequate energy resources, it would be impossible for economies to thrive and for people to lead comfortable lives. Therefore, it is essential to prioritize energy policies and investments to ensure a brighter future for all [19]. Sustainability is a crucial aspect of the global energy transition. All dimensions of sustainability must be taken into consideration while formulating and implementing policies, planning, operating, and dispatching energy resources for both generation and consumption. Only then can we ensure a sustainable future for generations [21]. As data flows between various components, providing robust encryption and intrusion detection systems is imperative to prevent potential cyber threats that could compromise the system’s reliability and security. For a long time, energy did not seriously factor in sustainable development. However,

sustainable development and sustainability issues now play a central role in energy and electricity by anchoring the evolution of the sustainable development paradigm. We must continue prioritizing sustainability in the energy sector to ensure a better future.

**6. Conclusion.** In conclusion, the stability study of a power system based on multi-intelligent body collaboration is a comprehensive assessment of this innovative approach's technical, operational, and economic aspects. It examines how the synergy between various intelligent entities can enhance grid stability, adaptability, and efficiency. By addressing challenges related to prediction accuracy, communication, resilience, and cost-effectiveness, this study paves the way for a more resilient and responsive energy infrastructure. The distributed nature of intelligent entities allows for real-time data collection and analysis, enabling proactive adjustments to prevent disruptions. This can lead to more efficient load management, reduced downtime, and enhanced grid stability. Moreover, the stability study would delve into these intelligent bodies' communication protocols and data-sharing mechanisms. Efficient data exchange is crucial for quick decision-making and coordinated actions. It's important to ensure that the communication network is robust, secure, and capable of handling the vast amounts of data generated by the system.

The future of power systems lies in the continued development and integration of multi-intelligent body collaboration. Future directions in this field include further advancements in artificial intelligence, the deployment of more sophisticated IoT devices, and the incorporation of advanced predictive analytics. These innovations will improve grid stability, adaptability, and sustainability. Additionally, research should focus on creating standardized protocols and frameworks that facilitate seamless interoperability among various intelligent agents. As the energy landscape evolves, power systems will need to adapt continuously, and multi-intelligent body collaboration will play a pivotal role in ensuring a reliable and efficient energy distribution network for the future.

**Fundings.** This work is funded by the Science and Technology Project of SGCC (Research on Wind Generator Transient Voltage Optimal Control Technology in Weak Power Grid) (4000-202214067A-1-1-ZN).

#### REFERENCES

- [1] A. ADEYEMI, M. YAN, M. SHAHIDEHPUR, C. BOTERO, A. V. GUERRA, N. GURUNG, L. C. ZHANG, AND A. PAASO, *Blockchain technology applications in power distribution systems*, The Electricity Journal, 33 (2020), p. 106817.
- [2] T. AHMAD, R. MADONSKI, D. ZHANG, C. HUANG, AND A. MUJEEB, *Data-driven probabilistic machine learning in sustainable smart energy/smart energy systems: Key developments, challenges, and future research opportunities in the context of smart grid paradigm*, Renewable and Sustainable Energy Reviews, 160 (2022), p. 112128.
- [3] F. E. ALIABADI, K. AGBOSSOU, S. KELOUWANI, N. HENAO, AND S. S. HOSSEINI, *Coordination of smart home energy management systems in neighborhood areas: A systematic review*, IEEE Access, 9 (2021), pp. 36417–36443.
- [4] T. DARWISH, G. K. KURT, H. YANIKOMEROGLU, G. SENARATH, AND P. ZHU, *A vision of self-evolving network management for future intelligent vertical hetnet*, IEEE Wireless Communications, 28 (2021), pp. 96–105.
- [5] D. FAQUIR, N. CHOULIARAS, V. SOFIA, K. OLGA, AND L. MAGLARAS, *Cybersecurity in smart grids, challenges and solutions*, AIMS Electronics and Electrical Engineering, 5 (2021), pp. 24–37.
- [6] Y. FENG ET AL., *Research of finite-time event-triggered control for high-speed train under switching topology*, Journal of Applied Science and Engineering, 26 (2022), pp. 933–940.
- [7] A. GOUDARZI, F. GHAYOOR, M. WASEEM, S. FAHAD, AND I. TRAORE, *A survey on iot-enabled smart grids: Emerging, applications, challenges, and outlook*, Energies, 15 (2022), p. 6984.
- [8] B. GUO, Y. LIU, S. LIU, Z. YU, AND X. ZHOU, *CrowdHMT: crowd intelligence with the deep fusion of human, machine, and iot*, IEEE Internet of Things Journal, 9 (2022), pp. 24822–24842.
- [9] J. HE, Y. LI, H. LI, H. TONG, Z. YUAN, X. YANG, AND W. HUANG, *Application of game theory in integrated energy system systems: a review*, Ieee Access, 8 (2020), pp. 93380–93397.
- [10] M. A. KHAN, N. KUMAR, S. A. H. MOHSAN, W. U. KHAN, M. M. NASRALLA, M. H. ALSHARIF, J. ŻYWOLEK, AND I. ULLAH, *Swarm of uavs for network management in 6g: A technical review*, IEEE Transactions on Network and Service Management, (2022).
- [11] J. LE, G. QI, L. ZHAO, AND R. JIN, *Research on a control strategy for a distributed economic dispatch system in an active distribution network, considering communication packet loss*, Electronics, 11 (2022), p. 3288.
- [12] S. LI, Y. JIA, F. YANG, Q. QIN, H. GAO, AND Y. ZHOU, *Collaborative decision-making method for multi-uav based on multiagent reinforcement learning*, IEEE Access, 10 (2022), pp. 91385–91396.
- [13] Y. LIU, X. LIU, X. MU, T. HOU, J. XU, M. DI RENZO, AND N. AL-DHAHIR, *Reconfigurable intelligent surfaces: Principles and opportunities*, IEEE communications surveys & tutorials, 23 (2021), pp. 1546–1577.

- [14] O. P. MAHELA, M. KHOSRAVY, N. GUPTA, B. KHAN, H. H. ALHELOU, R. MAHLA, N. PATEL, AND P. SIANO, *Comprehensive overview of multi-agent systems for controlling smart grids*, CSEE Journal of Power and Energy Systems, 8 (2020), pp. 115–131.
- [15] M. B. MOLLAH, J. ZHAO, D. NIYATO, Y. L. GUAN, C. YUEN, S. SUN, K.-Y. LAM, AND L. H. KOH, *Blockchain for the internet of vehicles towards intelligent transportation systems: A survey*, IEEE Internet of Things Journal, 8 (2020), pp. 4157–4185.
- [16] L. QIAN, P. YANG, M. XIAO, O. A. DOBRE, M. DI RENZO, J. LI, Z. HAN, Q. YI, AND J. ZHAO, *Distributed learning for wireless communications: Methods, applications and challenges*, IEEE Journal of Selected Topics in Signal Processing, 16 (2022), pp. 326–342.
- [17] M. H. SHAWON, S. MUYEEN, A. GHOSH, S. M. ISLAM, AND M. S. BAPTISTA, *Multi-agent systems in ict enabled smart grid: A status update on technology framework and applications*, Ieee Access, 7 (2019), pp. 97959–97973.
- [18] K. TAZI, F. M. ABBOU, AND F. ABDI, *Multi-agent system for microgrids: design, optimization and performance*, Artificial Intelligence Review, 53 (2020), pp. 1233–1292.
- [19] W. TUSHAR, C. YUEN, T. K. SAHA, T. MORSTYN, A. C. CHAPMAN, M. J. E. ALAM, S. HANIF, AND H. V. POOR, *Peer-to-peer energy systems for connected communities: A review of recent advances and emerging challenges*, Applied energy, 282 (2021), p. 116131.
- [20] X. WANG, J. LI, Z. NING, Q. SONG, L. GUO, S. GUO, AND M. S. OBAIDAT, *Wireless powered mobile edge computing networks: A survey*, ACM Computing Surveys, (2023).
- [21] P. ZHANG, H. XUE, AND S. GAO, *Multi-agent fault-tolerant control based on distributed adaptive consensus*, Ieee Access, 7 (2019), pp. 135882–135895.

*Edited by:* Sathishkumar V E

*Special issue on:* Scalability and Sustainability in Distributed Sensor Networks

*Received:* Aug 23, 2023

*Accepted:* Feb 13, 2024



## STUDY ON GRID-CONNECTED POWER QUALITY IMPROVEMENT OF WIND FARMS BASED ON REPETITIVE CONTROLLER

MINJIE ZHU\*, LIANGNIAN LV†, XUBO LE‡, AISIKAER§, HAIBO LI¶, AND YUCHENG GAO||

**Abstract.** The project highlights the wind energy and the process followed in the wind farms in order to generate the energy by using wind power and wind capacity as well. Moreover, it has been observed that wind energy is also considered one of the most effective electrical energy sources. The application of wind energy can be beneficial for different aspects of the business sectors of different countries across the globe and this is possible because of its cost-effectiveness. Furthermore, this study encompasses the strategies employed within wind farms to preserve and efficiently harness wind power for the generation of electrical energy. Notably, the project delineates the utilization of a diverse array of instruments, each instrumental in the conversion of wind power into valuable wind energy. Beyond the realm of technology, wind farms grapple with a spectrum of formidable challenges, stemming from the inherently capricious nature of wind. In addition, various issues concerning power quality and the stability of the power system have also been identified within these wind farms. To ameliorate these multifaceted challenges, the project introduces a range of meticulous control measures for implementation.

**Key words:** Wind farms, Wind turbines, motor, instruments, grid, access, power, issues, voltage.

**1. Introduction.** Wind farms are also known as wind power plants in which different modern technology-based equipment and software applications have been used to generate energy from the wind forces. It has been observed that the power of wind and the forces of wind are converted into electricity power. This kind of energy source which is provided by high-power wind sources is also known as renewable energy. The capacity of generating electricity-based energy is not limited to a particular amount. Moreover, this kind of energy source is the “lowest cost energy source”. Among different types of equipment, the most useful equipment utilized in wind farms in order to generate electricity from the wind forces is wind turbines. With the help of wind turbines, the developer can use the forces of surrounding winds and conserve them and convert them into electricity energy. Apart from different advantages it has been observed that different challenging situations are also observed in the wind farms while converting the wind forces into electricity energy. The uncontrolled nature of wind forces results in the development of different types of challenging situations. Apart from this various other issue are also observed in the wind farms such as power quality-related issues, and the stability of the power system-related issues as well.

As the global energy landscape undergoes a transformative shift towards sustainability and renewable sources, wind energy has emerged as a prominent and environmentally responsible solution. Wind farms, comprising arrays of wind turbines, play a pivotal role in the generation of clean electricity. However, the integration of wind energy into the electrical grid introduces a unique set of challenges, particularly with regard to power quality. Wind energy is inherently intermittent due to the variable nature of wind speeds. This intermittency can result in fluctuations in power generation and, consequently, issues related to power quality when integrated into the grid. Voltage and frequency variations, harmonics, and flicker are some of the power quality concerns that grid operators and utilities must contend with. The need to enhance the quality and reliability of electricity supplied to consumers underscores the significance of this research.

By conducting an in-depth analysis of the Repetitive Controller’s application in wind farms, this research

---

\*State Grid Taizhou Power Supply Company, China, 317700

†Goldwind Science & Technology Co., Ltd, Beijing, China, 100176

‡State Grid Taizhou Power Supply Company, China, 317700

§Goldwind Science & Technology Co., Ltd, Beijing, China, 100176 ([minjiezhu1@outlook.com](mailto:minjiezhu1@outlook.com))

¶State Grid Taizhou Power Supply Company, China, 317700

||Goldwind Science Technology Co., Ltd, Beijing, China, 100176

aims to contribute to the broader goal of optimizing the performance of wind energy systems while ensuring high-quality power delivery. It underscores the importance of stable and reliable electricity supplies as we strive to meet the ever-increasing demand for clean and sustainable energy. In a world seeking to reduce carbon emissions and mitigate the effects of climate change, the enhancement of grid-connected power quality in wind farms is a critical step towards realizing a more sustainable and environmentally responsible energy future.

**2. Objectives.** Different motivation has been observed behind the development of this study. Hence these motivations are further illustrated as the objectives of this study in the below section.

1. To identify the issues of different power quality control aspects of the wind farms
2. To investigate the challenging conditions the wind conservation farm can experience due to power grid
3. To find out the various ranges of controllers that are used in the wind farms
4. To describe the role of wind turbines in the enhancement of production capacity of the electricity energy in the wind farms
5. To examine the innovative ideas and their contribution to the improvement of the output level of wind farms

**3. Methodology.** The conducted study is based on the different information which is collected about wind farms and the technology used in wind farms as well. It has been observed that different types of motors and turbines are used in wind farms for converting wind energy into electricity power [1]. Information about different challenging conditions and the different innovative strategies adopted by wind farms are also incorporated in this study. It has been observed that all data are mainly secondary in nature. As the use of secondary information has been observed in the development of this research project, therefore, it can be stated that the use of the qualitative method is also present in the research project as well.

**3.1. Power Quality Issues of Wind Farms.** The motivation that has been observed behind the development of the wind farm is to utilize the wind forces and convert them into electric energy. The quantity of the electric energy can be described by its voltage and frequency [2]. Due to the variation in the frequency power generated in the form also varies as well. This can cause lower quality of the generated electric energy. Moreover, it has been observed that it is next to impossible to maintain a constant frequency and voltage for the electrical power generated by using the wind forces in the wind farm [3].

Wind energy, as an increasingly prevalent renewable energy source, has brought substantial benefits in reducing greenhouse gas emissions and enhancing sustainability. However, the integration of wind farms into the electrical grid introduces power quality challenges that must be effectively addressed. One of the primary issues pertains to the variable and intermittent nature of wind. Wind speeds fluctuate, leading to inconsistent power generation, which can result in voltage and frequency variations. These variations can, in turn, disrupt the stability of the grid and affect the quality of electricity supplied to consumers.

Harmonics are another significant concern associated with wind farms. Harmonics are unwanted, high-frequency electrical disturbances that can lead to equipment malfunction and efficiency losses. The non-linear characteristics of power electronic converters within wind turbines can introduce harmonics into the electrical grid, impacting the overall power quality. Additionally, flicker, which is the rapid variation in voltage magnitude, can be generated by wind turbines, causing undesirable fluctuations in lighting and equipment performance, and potentially leading to discomfort and inconvenience for end-users

Among different problems, the development of harmonic distortion is considered as one of the major issues, and due to this issue it has been observed that a wrong measurement has been detected by the wind turbine [4]. Among different reasons, some common reasons for developing this situation differ in the speed drivers of the wind farms. When this differentiation between two-speed drivers becomes then this causes the development of harmonic distortion. Moreover, it also needs to be noted that “large concentrations of arc discharge lamps”, and “saturation of magnetization of transformer” are another two reasons which also contrived in the development of the harmonic distortion [5]. Due to this harmonic distortion, the voltage in the grid of the wind farm got influenced and it started fluctuating as well. There are two different harmonic voltage grids that have been noticeable in the wind farm which affected the measurements of the voltage unit. These two grids are “harmonic of 5th and 7th grids”.

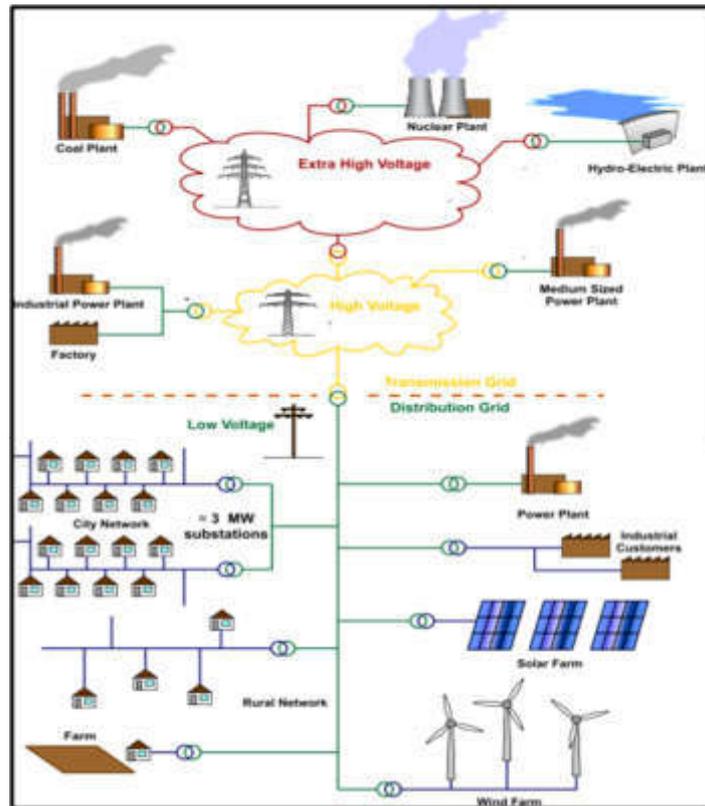


Fig. 3.1: Power quality issues of wind farms

**3.2. Power Grid Challenging Conditions of Wind Conservative Farm.** It has been observed that the “intermittent nature of the wind energy” is one of the highly influential challenging situations developed in the wind form which in turn creates power grid-related challenging conditions. In other power plants, the developer has the ability to control the power that is used in order to generate electrical energy [6]. Though in case it has been observed that the speed and the force of the air cannot be controlled by the developer as a result of this the developer finds it difficult to match a high-speed containing air and convert them into electrical energy as well. Due to this uncontrollable power of the air, the output generated after the conversion has also become out of control and out of expectation as well [7].

Moreover, the availability of wind is another challenging situation faced by the developers of wind farms. The high range of availability of the wind makes suitable conditions for generating electrical energy from the wind forces [8]. On the other hand, a minimum amount of wind availability is required for generating electrical energy. When the availability of the air is below the minimum range then this situation creates an inability in the development of electrical energy [9]. In addition to that it has been observed that it lowers the production rate of electrical energy. Moreover, the uncertain condition of the wind capacity also results in the development of different issues regarding the power quality issues of the produced electrical power. It has been observed that this uncertain condition also resulted in the development of instability in the electrical power as well [10].

**3.3. Types of Controllers used in Wind Farms.** Different types of controllers have been used in these wind farms and these controllers have been involved in impacting the production rate of the generated electrical energy [11]. Moreover, it has been observed that wind turbines are responsible for conserving the winds from the surrounding areas. Therefore, the winds collected by the wind turbines are monitored by the controller that is used in the wind farms as well. Along with different computer devices and OS, systems are also involved in

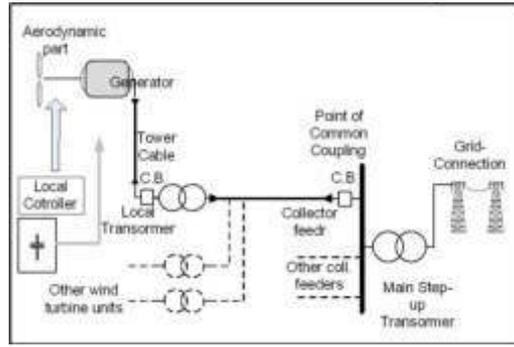


Fig. 3.2: Power grid challenging condition of wind conservative farm

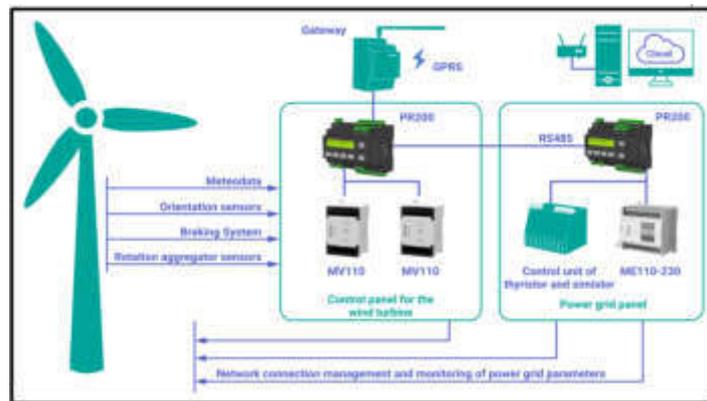


Fig. 3.3: Controllers used in wind farms

playing the role of the controller which can be able to control the speed and the capacity of the wind turbine that is used in the wind farms [12].

Apart from that the mostly used controllers of wind farms are “large numbers of switches”, “hydraulic pumps”, “valves, and different shaped motors” as well. These all have the ability to control the working capability of the wind turbulence that is used in wind farms [13]. Moreover, the size of the wind turbulence is also another aspect that can also control the power of the wind turbulence as well. These controlling machines are also helpful for impacting the performance capability of the wind turbulence and it enhances the capability of conserving the wind only in one cycle. When the wind conservation capacity becomes high then it is also helpful for enhancing the production rate of the generating electrical energy [14].

**3.4. Innovative Ideas that can improve the Wind Farm Output Level.** Different types of innovative ideas have been observed which are considered one of the measures that are helpful for enhancing the capability as well as the effectiveness of the wind farms in terms of maximizing the production rate of the electrical energy by those wind farms [15]. The inclusion of retrospective measures is considered as one of the innovative ideas and in case the developer used this measure in the controlling aspect then it has the capability to maximize the performance capability of the wind farms. Moreover, thus specific measures are also involved in developing different opportunities for controlling the capability of the wind turbines and maximizing the capacity of the wind turbines as well [17].

The measurement technology is also considered another measure of improving the output level of the wind farms [16]. By incorporating these measuring techniques the developer can be able to measure the availability of the wind and the spread of the wind in the surrounding areas of the wind farms and use these measurements

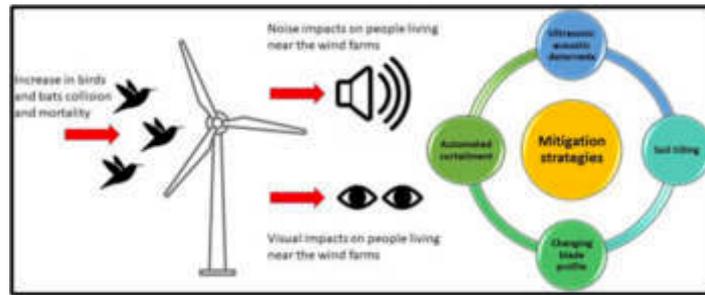


Fig. 3.4: Innovative ideas that can improve the wind farm output level

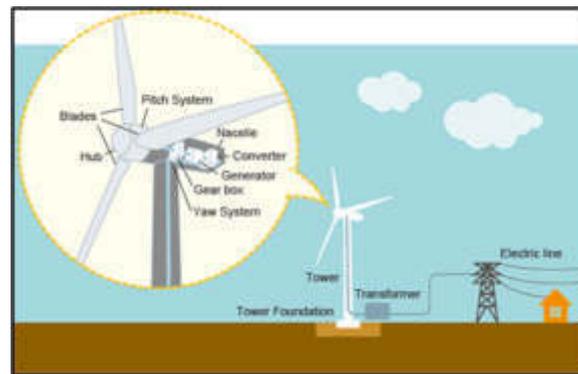


Fig. 3.5: Role of wind turbines

for further conserving the winds from the surrounding areas as well. Moreover, the application of advanced levels of Pitch, Yaw, and rotational speed is also helpful for enhancing the out level of the wind farms [17].

**3.5. Role of wind turbines in the wind farm.** Wind turbines are involved in contributing a high level of impact in the generation process of electrical energy [18]. Moreover, it has been observed that the application of wind turbines is involved in turning wind power into electrical energy. This whole process of turning wind power into electrical energy is also known as the process of generating electrical energy by using wind power and wind capacity as well. Moreover, it has been observed that the high range of the hands of wind turbines is helpful for arranging greater areas of the wind. With the long hand of the wind turbines, the reaching capacity of the wind turbines exceeds by a high level [19].

With the help of the inclusion of a high-speed-containing motor, the speed of the wind turbine can be increased and this impacts maximizing a greater amount of wind from the surrounding space [20].

**4. Results.** With the inclusion of different types of technology and software applications, it has been observed that the capacity of wind farms gets maximized and this is helpful for generating a high range of electrical energy. Moreover, the application of wind energy is increasing day by day in different countries due to its cost-effectiveness. The easy way of developing this renewable energy is also another reason for enhancing the usage of wind energy [21].

The above-concerned fig. 4.1 contains the percentage of the utilization of wind farms in different countries. It has been observed that in the year 2016, the UK does not have any exposure to wind farms. On the other hand, in the year 2022, the exposure of wind farms and their utilization in the generation of high levels of electrical energy has reached its maximum. Moreover, other countries of the world also reach their highest range in terms of using wind farms for different purposes of the development of the country.

The above-developed fig 4.2 contains the five different countries of the world which have the largest wind

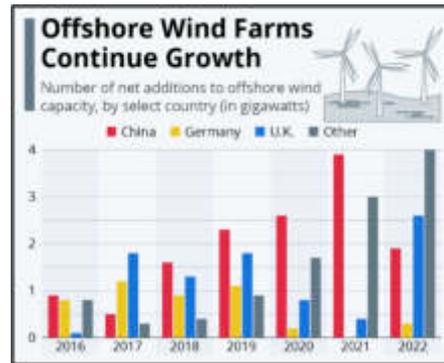


Fig. 4.1: Utilization of wind energy

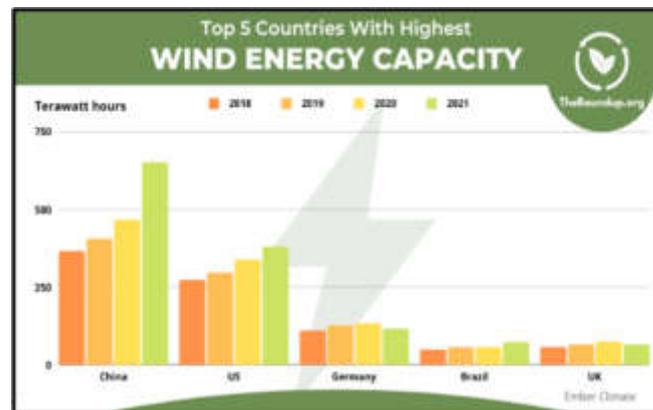


Fig. 4.2: Top 5 countries with wind energy capacity

farms and also contain the largest amount of wind energy capacity as well. Among the top 5 countries, China ranked in the first position, and in the year 2021, almost 750 terawatt hours of energy will be produced in China by using the wind forces in the wind farm as well.

**5. Conclusion.** By summarizing all the information thus it can be concluded that the concerned project contains all the information about the wind farms and the strategy used by the wind farms in order to conserve the wind and transform them in order to generate the electrical energy. It has been observed that different strategies, different technology, different instruments, and different techniques that are used for conserving the wind forces and further utilized for producing electrical energy are illustrated in this project. Moreover, different challenging situations faced by wind farms are also described as well. This challenging situation lowers the efficiency of the wind farms and as a result of this scenario, the production rate of generating the electrical energy becomes low. the study also highlights the challenges that wind farms face, which can diminish their efficiency and consequently impact the rate of electrical energy generation. Addressing these challenges and continuing to refine wind energy technologies are imperative steps in the journey towards a greener and more sustainable energy future. It underscores the global shift towards wind energy and emphasizes the importance of overcoming challenges to further enhance the efficiency and effectiveness of wind farm operations. As the world seeks to reduce carbon emissions and bolster its commitment to renewable energy, the role of wind farms in achieving these objectives remains pivotal.

## REFERENCES

- [1] O. ABEDINIA, M. LOTFI, M. BAGHERI, B. SOBHANI, M. SHAFIE-KHAH, AND J. P. CATALÃO, *Improved emd-based complex prediction model for wind power forecasting*, IEEE Transactions on Sustainable Energy, 11 (2020), pp. 2790–2802.
- [2] S. W. ALI, M. SADIQ, Y. TERRICHE, S. A. R. NAQVI, M. U. MUTARRAF, M. A. HASSAN, G. YANG, C.-L. SU, J. M. GUERRERO, ET AL., *Offshore wind farm-grid integration: A review on infrastructure, challenges, and grid solutions*, IEEE Access, 9 (2021), pp. 102811–102827.
- [3] A. A. ALSAKATI, C. A. VAITHILINGAM, AND J. ALNASSEIR, *Transient stability assessment of ieee 9-bus system integrated wind farm*, in MATEC Web of Conferences, vol. 335, EDP Sciences, 2021, p. 02006.
- [4] S. ASADOLLAH, R. ZHU, AND M. LISERRE, *Analysis of voltage control strategies for wind farms*, IEEE Transactions on Sustainable Energy, 11 (2019), pp. 1002–1012.
- [5] W. BAO, Q. WU, L. DING, S. HUANG, AND V. TERZIJA, *A hierarchical inertial control scheme for multiple wind farms with besss based on adm*, IEEE Transactions on Sustainable Energy, 12 (2020), pp. 751–760.
- [6] H.-M. CHUNG, S. MAHARJAN, Y. ZHANG, F. ELIASSEN, AND K. STRUNZ, *Placement and routing optimization for automated inspection with unmanned aerial vehicles: A study in offshore wind farm*, IEEE Transactions on Industrial Informatics, 17 (2020), pp. 3032–3043.
- [7] S. DAWN, S. GOPE, S. S. DAS, AND T. S. USTUN, *Social welfare maximization of competitive congested power market considering wind farm and pumped hydroelectric storage system*, Electronics, 10 (2021), p. 2611.
- [8] S. I. EVANGELINE AND P. RATHIKA, *Wind farm incorporated optimal power flow solutions through multi-objective horse herd optimization with a novel constraint handling technique*, Expert Systems with Applications, 194 (2022), p. 116544.
- [9] S. HADAVI, M. Z. MANSOUR, AND B. BAHRANI, *Optimal allocation and sizing of synchronous condensers in weak grids with increased penetration of wind and solar farms*, IEEE Journal on Emerging and Selected Topics in Circuits and Systems, 11 (2021), pp. 199–209.
- [10] M. KHESHTI, L. DING, W. BAO, M. YIN, Q. WU, AND V. TERZIJA, *Toward intelligent inertial frequency participation of wind farms for the grid frequency control*, IEEE Transactions on Industrial Informatics, 16 (2019), pp. 6772–6786.
- [11] E. NADERI, M. POURAKBARI-KASMAEI, AND M. LEHTONEN, *Transmission expansion planning integrated with wind farms: A review, comparative study, and a novel profound search approach*, International Journal of Electrical Power & Energy Systems, 115 (2020), p. 105460.
- [12] J. QI, W. ZHAO, AND X. BIAN, *Comparative study of svc and statcom reactive power compensation for prosumer microgrids with dfig-based wind farm integration*, IEEE Access, 8 (2020), pp. 209878–209885.
- [13] B. SHAO, S. ZHAO, Y. YANG, B. GAO, L. WANG, AND F. BLAABJERG, *Nonlinear subsynchronous oscillation damping controller for direct-drive wind farms with vsc-hvdc systems*, IEEE Journal of Emerging and Selected Topics in Power Electronics, 10 (2020), pp. 2842–2858.
- [14] N. R. SULAKE, S. PRANUPA, AND A. SRIRAM, *A review of wind farm layout optimization techniques for optimal placement of wind turbines*, International Journal of Renewable Energy Research (IJRER), 13 (2023), pp. 957–965.
- [15] S. TAO, S. KUENZEL, Q. XU, AND Z. CHEN, *Optimal micro-siting of wind turbines in an offshore wind farm using frandsen-gaussian wake model*, IEEE Transactions on Power Systems, 34 (2019), pp. 4944–4954.
- [16] S. TAO, Q. XU, A. FEIJÓO, P. HOU, AND G. ZHENG, *Bi-hierarchy optimization of a wind farm considering environmental impact*, IEEE Transactions on Sustainable Energy, 11 (2020), pp. 2515–2524.
- [17] S. TAO, Q. XU, A. FEIJÓO, AND G. ZHENG, *Joint optimization of wind turbine micrositing and cabling in an offshore wind farm*, IEEE Transactions on Smart Grid, 12 (2020), pp. 834–844.
- [18] T. WANG, S. HUANG, M. GAO, AND Z. WANG, *Adaptive extended kalman filter based dynamic equivalent method of pmsg wind farm cluster*, IEEE Transactions on Industry Applications, 57 (2021), pp. 2908–2917.
- [19] Y. ZHANG, Y. YANG, X. CHEN, AND C. GONG, *Intelligent parameter design-based impedance optimization of statcom to mitigate resonance in wind farms*, IEEE Journal of Emerging and Selected Topics in Power Electronics, 9 (2020), pp. 3201–3215.
- [20] S. ZHENG, D. SONG, M. SU, X. YANG, Y. H. JOO, ET AL., *Comprehensive optimization for fatigue loads of wind turbines in complex-terrain wind farms*, IEEE Transactions on Sustainable Energy, 12 (2020), pp. 909–919.
- [21] Y. ZHOU, L. ZHAO, I. B. MATSUO, AND W.-J. LEE, *A dynamic weighted aggregation equivalent modeling approach for the dfig wind farm considering the weibull distribution for fault analysis*, IEEE Transactions on Industry Applications, 55 (2019), pp. 5514–5523.

*Edited by:* Sathishkumar V E

*Special issue on:* Scalability and Sustainability in Distributed Sensor Networks

*Received:* Aug 24, 2023

*Accepted:* Oct 29, 2023



## SMART FARMING USING THE BIG DATA-DRIVEN APPROACH FOR SUSTAINABLE AGRICULTURE WITH IOT-DEEP LEARNING TECHNIQUES

BUYU WANG\*

**Abstract.** The study showcases the process of deep learning operated in agriculture, including Deep IoT, which makes the procedure easier using the deep neural system. The use of the IoT in the agrarian sectors makes the evolution of firms more effective. The application of the IoT detector supports the making of grade derivatives in the husbandry department. Marketing of crop finance is two other operations of smart agriculture that help for better harvest farming. Through the IoT technology in the farming industry, agriculturalists can get notifications about the temperature and climate. The method needs professional and qualified employees in the division to properly monitor the system and the methods.

The submission of the proper nourishment for the proper crop increases the life duration of the harvest and makes the crop free from menace. The velocity of the manufacture of undeveloped items can also be improved by using the IoT. The function of the BDA and IoT has enlarged for the healthier construction of farming items. The foreword of elegant farming in the rural industry requires more capable and qualified trainers to give the personnel proper teaching. The urbanization of the farming process and the use of elegant and modern technology are well-designed in the time of “Agriculture 3.0.”

**Key words:** Big-data-driven approach, IoT, smart farming, Sensor technologies, deep learning, Agriculture etc.

**1. Introduction.** The big-data-driven approach or BDA is a process by which a decision is taken with the help of the observation of the data by the complex data. With the help of the IoT, the improvement in agriculture has increased [22]. The algorithm process in deep learning used in agriculture includes DeepIoT, which makes the process more accessible using the deep neural system. The IoT helps to connect the devices to the internet, which is helpful for the better performance of the system. UAV used in the IoT technology is one of the most effective for the agricultural department.

The main advantageous of smart farming are as follows,

1. Automation of tasks like planting, harvesting, and irrigation reduces the need for manual labor, resulting in increased efficiency and reduced labor costs.
2. Precision application of fertilizers, pesticides, and water ensures that resources are used more efficiently.
3. Smart farming technologies enable farmers to monitor and control various environmental factors such as soil moisture, temperature, and nutrient levels, leading to better crop management and higher yields.
4. Early disease detection and pest control help protect crops, ensuring a more abundant harvest.
5. Reduced use of resources such as water and chemicals lead to a more sustainable and environmentally friendly farming approach.

Smart farming, also known as precision agriculture, is an innovative approach to farming that leverages technology to enhance productivity, efficiency, and sustainability in agriculture. It involves the use of various technologies such as sensors, automation, data analytics, and connectivity to optimize various aspects of farming.

### 2. Objectives.

1. To understand the approach of BDA in smart farming
2. To find the feasible techniques of BDA that are helpful for the agricultural department
3. To determine the issues of the use of the IoT process in smart farming
4. To estimate the future effectiveness of deep learning in the agricultural department
5. To assess the effectiveness and impact of the use of modern technology in the farming department

The main motivation behind the research is ,

---

\*College of Computer and Information Engineering, Inner Mongolia Agricultural University, Hohhot 010018, China

1. The primary motivation is to improve the efficiency and effectiveness of agriculture. The use of deep learning and IoT technologies in agriculture is seen as a way to streamline various processes and make them more efficient, leading to better outcomes in terms of crop yield, quality, and resource utilization.
2. The research aims to explore the application of the Internet of Things (IoT) in the agricultural sector. IoT can offer real-time monitoring and data collection capabilities, which are vital for precision agriculture. The motivation is to leverage IoT technology to enhance decision-making and overall farm management.
3. The study seeks to contribute to the evolution of agricultural practices and the agricultural industry as a whole. It recognizes that modern technology, such as IoT and deep learning, can lead to more advanced and effective farming methods and help agricultural firms progress.

**3. Methodology.** The use of a vast number of data and information used in the study is helpful for the analysis of the impact of the IoT in smart farming. The data and information are collected from various kinds of secondary sources like the internet, journals, and articles. The application of the IoT in the agricultural department makes the growth of firms more convenient [13]. The secondary method used in the study is beneficial for analyzing the impact of deep learning in smart farming.

**3.1. The approach of BDA in smart farming.** Different approaches and techniques are used in the BDA, which are effective for smart farming in the present day. The application of the water cycles, the rainfall patterns and the fertilizer requirements are the most effective and valuable for smart learning [26]. Smart farming is the process of farming that uses advanced techniques and IoT technologies for the better growth and development of agricultural items. Applying artificial intelligence techniques is also a helpful tool for the smart farming process.

**3.1.1. The approach used in the intelligent farming.** There are many kinds of factors and techniques that are used in smart farming for the better quality of the products. *Sensor technologies* are crucial techniques for managing the temperature and other natural factors used in intelligent farming [12]. The ordinance of the IoT sensor helps the making of quality products in the agriculture department.

The process of integration of the smart technique is the main factors of the implementation of the new and innovative technique in the smart farming. The upgradation of the traditional methods and enhancement of the quality and production of agricultural products is another factor of the smart farming technology. The communication process and use of innovative technology is the one of the most helpful approaches for the development of the process of smart farming [11]. There are various kinds of methods and techniques of the smart farming like precession farming, aquaponics, hydroponics and the vertical farming.

The use of the different kinds of techniques is different and the application of this techniques increase the efficiency of the methods and also reduce the workload of the farmers. The popularity of the above techniques and methods gaining increasingly day by day among the farmers. The application of this techniques in the farming sector also reduces the effort of the workers which leads to the increment of the rate of requirement of the employees in the business.

From the figure 3.1, it can be said that the core process of the smart operation of smart farming is the storage of the data in the data cloud and the proper analysis of the data and information [28]. The constellation of the process and the planning of the farming of the crop have to be done by the proper verification by the IoT technologies [10]. Marketing and the insurance of crop finance are two other operations of smart farming that help for better farming of the crop [17]. Therefore, the proper implementation of the techniques and avoidance of the traditional practices and processes become helpful for the enhancement of smart farming in the present day.

The figure 3.2 shows that the yield prediction and the prediction of the weather affect the growth of the farming of agricultural items. The processing of the data and the analysis of the information help to increase the production of farming items [23]. The analysis of the data related to the quality of soil and the weather are two effectives for better farming. The improvement of the communication process is smart farming and the recruitment of skilled workers helps to implement the process in the growth of the agricultural items in the farming industries [2]. Therefore, the techniques and approaches that are used in the BDA for smart farming are useful for enhancing sustainability in the agricultural department.

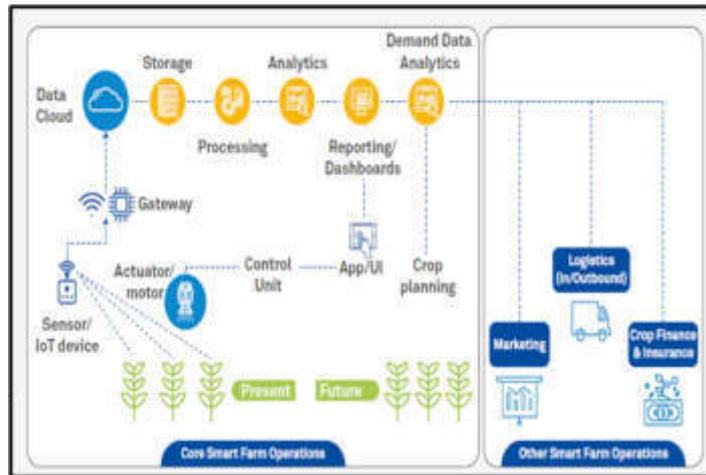


Fig. 3.1: The process of smart farming

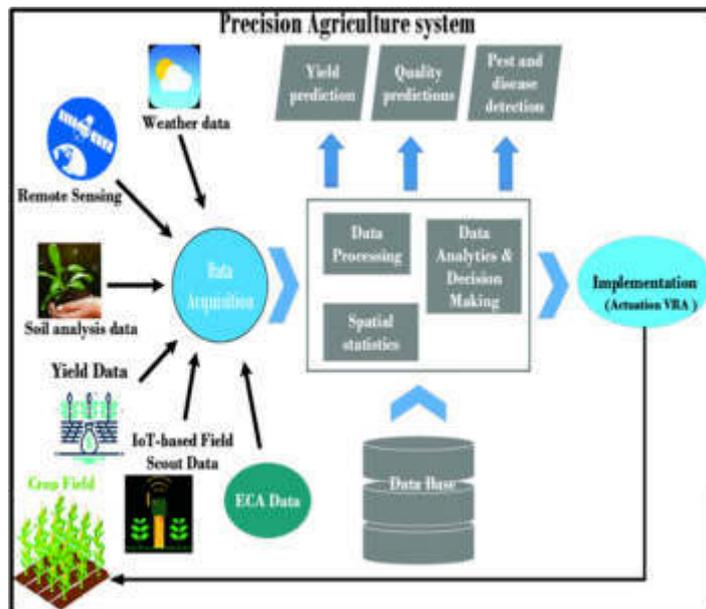


Fig. 3.2: Representation of the big data in smart farming

**3.2. Techniques of BDA for the agricultural department.** Different kinds of techniques and process are used in the agricultural industries for making a revolution in the industry. From the year 1890 to the year 2020, the agricultural department faced various kinds of revolution in the department. The implementation of the machine in the agricultural department was shown from the year 1890 to the year 1940 [24]. For farming, the proper knowledge of the climate and weather is very important to increase the rate of farming in a particular area [1]. With the help of IoT technology in the farming industry, farmers can be able to get information about the weather and climate. The use of the smart farming sensor becomes beneficial for knowing the condition of the weather.

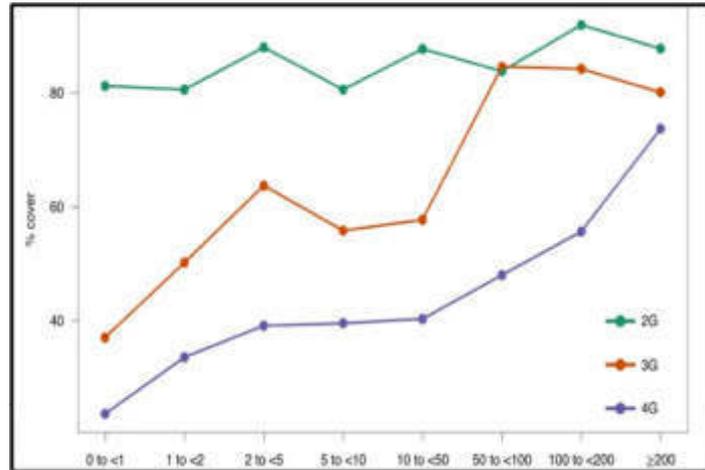


Fig. 3.3: Data-driven farming in the world

**3.2.1. Livestock monitoring.** Enabling the livestock monitoring process by using the tag or collar on the animals, the health condition, location and blood pressure and many other things of the animal can be detected. This device and the process is helpful in the agriculture sector and helps the farmers to understand the unhealthy things that is out of the range of normal factors. With the help of the livestock monitoring process the farmers can be able to understand the viral illness of the animal and after the detection they can be able to take the proper step according to the level. Other than the detection of the health of the animal, the monitoring system is also affective for the tracking of the GPS for collecting the historic data and information about the grazing place.

The main benefits and effectiveness of the livestock monitoring process are:

1. The analysing of the time to give birth of the animals and determining the breeding
2. Tracking of the patterns of grazing and the appropriate place for grazing
3. Determining the condition of health of the animals
4. Analysing the data and information that are historic

With the help of smart technologies like IoT technologies and the BDA, the process of smart farming becomes easier. Through the process of livestock monitoring, the health of the animals can be determined [21]. The detention of the diseases of the animals and the process of recovery of the animal from the diseases can be easily estimated by the above process.

**3.2.2. Greenhouse automation.** One of the most impactful and effective technologies in smart farming is the process of the greenhouse automation. With the use of the BDA and IoT technologies, the detection of the level of greenhouse gases becomes easier for farmers [5]. The future detection of the condition of the soil and the level of humidity can be estimated with the help of this process. In the process of BDA, the detection of the weather using the smart weather station is also useful to understand the condition of the weather for better farming.

The figure 3.3 shows the range of the usage of the data-driven process globally. Most of the countries in the world use the 2G data for smart farming and the range of 80% of the cover of the data is from 100 to 200. The usage of the 4G data is in the range below 40% and for the range greater than 200, 60% of the 4G data is used. The use of the various new and innovative technologies in the agricultural sector helps to err the effect of the greenhouse gases in the present days. The data and information that are used in this sector are collected from the internet and the authentic sources helps to emission the effect of the greenhouse and harmful gases. The technique of the data-driven in the agricultural sector helps to automate the harmful impact of the harmful emission. The installation of the software technologies and the automated actions helps to decrease the temperature in the environment with the opening of the process of ventilation. Thus, the controlling of the

temperature under the greenhouse automation is useful for the improvement of the works in the agriculture department.

**4. Issue of the IoT process in smart farming.** The use of the IoT in smart farming also includes some issues and problems in agriculture. The common issue with the use of the IoT in smart farming is the lack of skilled and knowledgeable workers in the IoT [14]. The scarcity of skilled workers in this sector made a negative impact on the application of various kinds of methods and technologies for better work in the process of smart farming [25]. The IoT system in smart farming is mainly used for the monitoring of the use of the sensor and the analysis of the automotive irrigation system [20]. Thus, the process needs more experienced and skilled employees in the department for the proper monitoring of the system and the technologies.

Among the all issues, the most vital and effective issues of the use of the IoT farming in the agricultural sector is the problems in the assess in the internet process. The lack of the internet connection in the rural sector makes the process of the agriculture less effective. In the year 2022, the value of the IoT in the agricultural sector is 13.7 US dollars and the value will be increased to 28.56 US dollars in the year 2030. The excessive use of the IoT things in the agricultural sector required the high skilled employees in the agricultural firm. Therefore, the lack in the number of the workers in the agricultural sector is a big issue of this department.

One of the most vital issues of the use of the IoT in smart farming is the location of the firms. Most of the agricultural firms are located in remote areas and rural areas. For rural areas, the *scarcity of the internet* and the issues of the network convention become the main problems for the agricultural department [3]. Also, the tools and the equipment of the IoT are of *high cost* and this is one of the most vital problems for smart farming [11]. The need for the proper resources of money makes the use and the process of the IoT less effective for the implementation of smart farming in the present day [7]. Another challenge that is faced by the agricultural sector is the lack of security for the employees and farmers in the industries. The increment of insecurity among the farmers is one of the big issues for smart farming. The implementation of the technology and modern internet process requires more security for the agricultural department and this helps in the growth and development of smart farming in the recent era.

**5. The future effectiveness of deep learning in agriculture.** There are both positive and negative effects of the use of the BDA and IoT in the smart farming. The application of the IoT in the farming industry becomes helpful and effective if the techniques can be used wisely [29]. With the help of the deep learning process, the workers in the agricultural department can be estimated the crop diseases and also estimate the benefits of the fertilizer [4]. The application of the appropriate fertilizer for the appropriate crop increases the life span of the crop and makes the crop free from danger [15]. The reduction of the cost can be happened by the use of deep learning in the agricultural department. Therefore, by using the process of deep learning, the cost and the price of the crops and other agricultural items can be minimized.

The IoT has a great future in the agricultural department and the use of this technology will be helpful for the better benefit of the farmers. The use of the IoT in the agricultural department helps to optimize the resources of the crops [30]. The rate of the production of agricultural items also can be enhanced by the use of the IoT. Monitoring the climate and the farming of the greenhouse become easier with the use of technology in the farming industries.

The figure 5.1 shows that in the year 2021, the rate of precession farming is the highest among all other techniques. In the year 2030, the use of techniques that are effective for smart farming becomes increased [27]. Therefore, it is clear that in the future, the application of the BDA and IoT become increased for the better production of agricultural items.

**6. Results.** It is clear from the above discussion that the application of the various kinds of technologies and methods in smart farming makes the process of farming more interesting and easier. Deep learning is one of the most impactful and effective ways to increase the rate of the production of agricultural items [16]. As a result, the introduction of smart farming in the agricultural industry requires more skilled and experienced trainers for giving the workers proper training [8]. The implementation of IoT technologies and the BDA in the agricultural department needs more money and resources [9]. The scarcity of money and resources become one of the most common issues in the agricultural department.

The figure 6.1 shows the aspects of the IoT in the Department of Agriculture. The strategic feature of

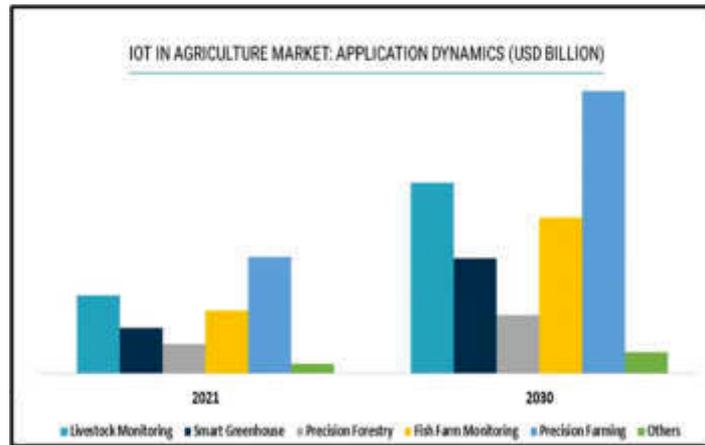


Fig. 5.1: Future of IoT in the agricultural department

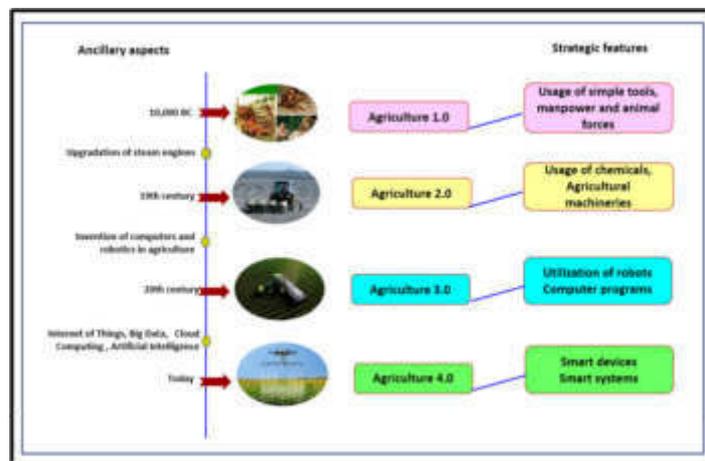


Fig. 6.1: Smart Farming and the IoT

“*Agriculture 1.0*” is the application of tools that are simple and easy to use [18]. The manpower and the use of animals are the main processes of this kind of agriculture. In the part of “*Agriculture 2.0*”, there was used various kinds of machines and chemicals. This upgradation process increases the rate of productivity of the agricultural department in the past years [19].

The figure 6.2 shows ways smart farming is shaping the growth of farming style through the usage of technology. The urbanization of the agricultural process and the use of smart and modern technologies are applied in the time of “*Agriculture 3.0*” The last stage of the agricultural process, the process of “*Agriculture 4.0*” makes the smart devices and the smart process for the growth and development of the sector [6]. From the above research it can be conclude that the innovative technologies in the data science gives a new revolution for the development and growth of the agricultural sector. The skilled based agriculture in the present days is the important factor that driven the technology and information sector to a high level. The emerging of the countries by the revolution of the agriculture department through the world improve the entire cycle of the agricultural sector. Thus, the use of the BDA and IoT technologies in smart farming is useful for the increment of the rate of production and the quality of the crops.

The initiatives of the government and NGO helps to driven the agricultural department to a next level. As

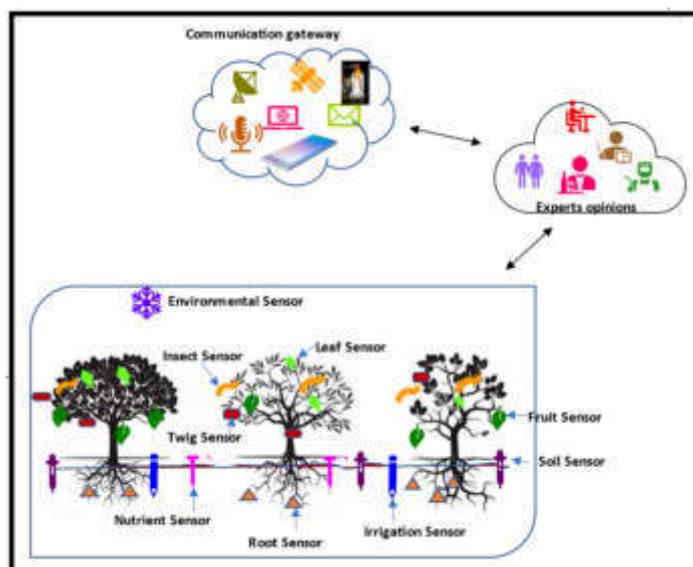


Fig. 6.2: Smart Farming

a result, it can be said that the changing of the weather is the one of the most effective for the development of the agricultural department and the growth in the production rate. The growth of the plants and crops can be detected properly with the help of the detection of the temperature of soil and weather. The implementation of the new technologies and methods in the agricultural sectors is the most useful and valuable process in the present days.

**7. Conclusion.** In conclusion, the study underscores the substantial influence of Internet of Things (IoT) and Big Data Analytics (BDA) on the field of agriculture. It demonstrates that the adoption of IoT and BDA technologies has brought about a profound transformation in agriculture, marking a shift from the traditional "Agriculture 1.0" to the more advanced and data-driven "Agriculture 4.0." This transition reflects the effective application of innovative methods and technologies, resulting in increased efficiency, sustainability, and productivity within the agricultural sector. Nevertheless, the conclusion also acknowledges that the integration of modern technologies is not without its challenges. One such challenge is the need for careful planning and the recruitment of skilled and capable workers in the agricultural industry to effectively harness the potential of IoT and BDA. Looking forward, the conclusion anticipates further advancements in agricultural production and crop quality, as the continued utilization of IoT in smart farming promises to bring about noticeable positive changes in the agricultural landscape. This future-oriented perspective highlights the ongoing opportunities for growth and development in agriculture through the incorporation of IoT technology.

## REFERENCES

- [1] M. S. ALI, A. M. M. RAAFEI, AND A. S. M. MIA, *Smart farming using ai towards bangladeshi agriculture*, Journal of FST, 1 (2022), p. 189.
- [2] M. AMIRI-ZARANDI, R. A. DARA, E. DUNCAN, AND E. D. FRASER, *Big data privacy in smart farming: a review*, Sustainability, 14 (2022), p. 9120.
- [3] M. AMIRI-ZARANDI, M. HAZRATI FARD, S. YOUSEFINAGHANI, M. KAVIANI, AND R. DARA, *A platform approach to smart farm information processing*, Agriculture, 12 (2022), p. 838.
- [4] M. ANDRONIE, G. LĂZĂROIU, M. IATAGAN, I. HURLOIU, AND I. DIJĂRESCU, *Sustainable cyber-physical production systems in big data-driven smart urban economy: a systematic literature review*, Sustainability, 13 (2021), p. 751.
- [5] S. A. BHAT AND N.-F. HUANG, *Big data and ai revolution in precision agriculture: Survey and challenges*, IEEE Access, 9 (2021), pp. 110209–110222.
- [6] N. CHANDRA, *Significant advancements in agriculture science*, Dr. Ajay B. Jadhao, (2023), p. 72.

- [7] J. A. DELGADO, N. M. SHORT JR, D. P. ROBERTS, AND B. VANDENBERG, *Big data analysis for sustainable agriculture on a geospatial cloud framework*, *Frontiers in Sustainable Food Systems*, 3 (2019), p. 54.
- [8] K. DINEVA AND T. ATANASOVA, *Cloud data-driven intelligent monitoring system for interactive smart farming*, *Sensors*, 22 (2022), p. 6566.
- [9] F. N. FOTE, S. MAHMOUDI, A. ROUKH, AND S. A. MAHMOUDI, *Big data storage and analysis for smart farming*, in 2020 5th International Conference on Cloud Computing and Artificial Intelligence: Technologies and Applications (CloudTech), IEEE, 2020, pp. 1–8.
- [10] P. FRAGA-LAMAS, L. RAMOS, V. MONDÉJAR-GUERRA, AND T. M. FERNÁNDEZ-CARAMÉS, *A review on iot deep learning uav systems for autonomous obstacle detection and collision avoidance*, *Remote Sensing*, 11 (2019), p. 2144.
- [11] M. GUPTA, M. ABDELSALAM, S. KHORSANDROO, AND S. MITTAL, *Security and privacy in smart farming: Challenges and opportunities*, *IEEE access*, 8 (2020), pp. 34564–34584.
- [12] B. GÜVEN, İ. BAZ, B. KOCAOĞLU, E. TOPRAK, D. EROL BARKANA, AND B. SOĞUTMAZ ÖZDEMİR, *Smart farming technologies for sustainable agriculture: From food to energy*, in *A Sustainable Green Future: Perspectives on Energy, Economy, Industry, Cities and Environment*, Springer, 2023, pp. 481–506.
- [13] M. JAVAID, A. HALEEM, R. P. SINGH, AND R. SUMAN, *Enhancing smart farming through the applications of agriculture 4.0 technologies*, *International Journal of Intelligent Networks*, 3 (2022), pp. 150–164.
- [14] E. KARUNATHILAKE, A. T. LE, S. HEO, Y. S. CHUNG, AND S. MANSOOR, *The path to smart farming: Innovations and opportunities in precision agriculture*, *Agriculture*, 13 (2023), p. 1593.
- [15] A. LYTOS, T. LAGKAS, P. SARIGIANNIDIS, M. ZERVAKIS, AND G. LIVANOS, *Towards smart farming: Systems, frameworks and exploitation of multiple sources*, *Computer Networks*, 172 (2020), p. 107147.
- [16] E. MICHENI, J. MACHII, AND J. MURUMBA, *Internet of things, big data analytics, and deep learning for sustainable precision agriculture*, in 2022 IST-Africa Conference (IST-Africa), IEEE, 2022, pp. 1–12.
- [17] G. OBI REDDY, B. DWIVEDI, AND G. RAVINDRA CHARY, *Applications of geospatial and big data technologies in smart farming*, in *Smart Agriculture for Developing Nations: Status, Perspectives and Challenges*, Springer, 2023, pp. 15–31.
- [18] K. PAUL, S. S. CHATTERJEE, P. PAI, A. VARSHNEY, S. JUIKAR, V. PRASAD, B. BHADRA, AND S. DASGUPTA, *Viable smart sensors and their application in data driven agriculture*, *Computers and Electronics in Agriculture*, 198 (2022), p. 107096.
- [19] A. REJEB, K. REJEB, AND S. ZAILANI, *Big data for sustainable agri-food supply chains: a review and future research perspectives*, *Journal of Data, Information and Management*, 3 (2021), pp. 167–182.
- [20] S. ROGOTIS AND N. MARIANOS, *Smart farming for sustainable agricultural production*, *Big Data in Bioeconomy: Results from the European DataBio Project*, (2021), pp. 191–205.
- [21] A. ROUKH, F. N. FOTE, S. A. MAHMOUDI, AND S. MAHMOUDI, *Big data processing architecture for smart farming*, *Procedia Computer Science*, 177 (2020), pp. 78–85.
- [22] R. SAADANE, A. CHEHRI, S. JEON, ET AL., *AI-based modeling and data-driven evaluation for smart farming-oriented big data architecture using iot with energy harvesting capabilities*, *Sustainable Energy Technologies and Assessments*, 52 (2022), p. 102093.
- [23] M. N. SADIKU, T. J. ASHAOLU, AND S. M. MUSA, *Big data in agriculture*, *Int. J. Sci. Adv*, 1 (2020), pp. 44–48.
- [24] V. SAIZ-RUBIO AND F. ROVIRA-MÁS, *From smart farming towards agriculture 5.0: A review on crop data management*, *Agronomy*, 10 (2020), p. 207.
- [25] M. N. I. SARKER, M. S. ISLAM, H. MURMU, AND E. ROZARIO, *Role of big data on digital farming*, *Int. J. Sci. Technol. Res*, 9 (2020), pp. 1222–1225.
- [26] M. N. I. SARKER, M. WU, B. CHANTHAMITH, S. YUSUFZADA, D. LI, AND J. ZHANG, *Big data driven smart agriculture: Pathway for sustainable development*, in 2019 2nd international conference on artificial intelligence and big data (ICAIBD), IEEE, 2019, pp. 60–65.
- [27] T. A. SHAIKH, W. A. MIR, T. RASOOL, AND S. SOFI, *Machine learning for smart agriculture and precision farming: towards making the fields talk*, *Archives of Computational Methods in Engineering*, 29 (2022), pp. 4557–4597.
- [28] A. SINGH, R. MEHROTRA, V. D. RAJPUT, P. DMITRIEV, A. K. SINGH, P. KUMAR, R. S. TOMAR, O. SINGH, AND A. K. SINGH, *Geoinformatics, artificial intelligence, sensor technology, big data: emerging modern tools for sustainable agriculture*, *Sustainable Agriculture Systems and Technologies*, (2022), pp. 295–313.
- [29] H. SINGH, N. HALDER, B. SINGH, J. SINGH, S. SHARMA, AND Y. SHACHAM-DIAMAND, *Smart farming revolution: portable and real-time soil nitrogen and phosphorus monitoring for sustainable agriculture*, *Sensors*, 23 (2023), p. 5914.
- [30] S. YADAV, A. KAUSHIK, M. SHARMA, AND S. SHARMA, *Disruptive technologies in smart farming: an expanded view with sentiment analysis*, *AgriEngineering*, 4 (2022), pp. 424–460.

*Edited by:* Sathishkumar V E

*Special issue on:* Scalability and Sustainability in Distributed Sensor Networks

*Received:* Aug 25, 2023

*Accepted:* Oct 30, 2023





## SCALABLE SOLUTIONS FOR WIND AND SOLAR DISTRIBUTED GENERATION: A STUDY OF PARALLEL ALGORITHMS IN A SMART GRID ENVIRONMENT

JIANZHONG LI, FEIPING YANG, YU CHEN, LI AO, AND WEI XIAO\*

**Abstract.** The main application of the informative data in the sector which are related to the energy are defines and explains as one of the crucial elements of Energy Internet. The advancement of the grid system are very vital as well as promising and faces many issues that are connected with the implementation of the renewable energy including solar and wind energy. The capacity of collecting of the data is the main elements of make are easy in taking decisions. The advancement of the technologies and its improvement has many benefits and advantages which was shown by the data analytic of the renewable source of energy in the various power stations. This is the framework which shows the development and growth of the potential establishment of the analyzation of the data and information in the smart grid and the utilities of power by the renewable resources. The seven domains and approaches are used for the purposed of predicting the stability, flexibility and safety from the advancement of the grid system. The secondary qualitative methods are used to define and explain the importance of the grid system in relation with the renewable source of energy that is wing and solar energy.

**Key words:** Grid system, renewable source of energy, Energy elements, seven domains etc.

**1. Introduction.** The term scalability or scalable solutions is referred to or defined as the capacity or the capability of a network or the system which are being developed or transformed easily [3]. Its main aim is to fulfill or satisfy the growing or increasing demands and wants for enhancing or developing the smart grid which is accepted to be highly desirable in increasing the scalability [15]. The term distributed generation is also connected to scalability or the scalable solution. The distributed generations are basically defined or described as a variety of technologies that are used in generating electricity near or at the places where are been used, that includes, winds, solar panels, and a combination or mixture of heat and power [14]. The smart grid environment in connection with the solar and wind are basically a system of digitalization that mostly consists of controlling, monitoring, analyzing, and examining the flow of bidirectional power. The term bidirectional power flow includes various generating units like hydropower generation, thermal power stations, nuclear power stations, end-use of consumers, and solar power plants [10]. In short, the scalable solution is defined as the techniques and approaches or the system model that works or functions and describes its ability to perform or cope under an increased workload.

The transition to clean and sustainable energy sources is a pressing global concern, and wind and solar distributed generation have emerged as key players in this paradigm shift. As the demand for renewable energy continues to grow, there is a critical need for scalable solutions that can efficiently harness and manage power from these sources [3]. In this context, this paper explores innovative strategies and technologies designed to facilitate the scalability of wind and solar distributed generation systems. By addressing the challenges and opportunities inherent in this field, this research aims to pave the way for a more sustainable and scalable energy future.

Renewable wind energy is a vital component of the modern energy landscape, offering a clean, sustainable, and increasingly important source of power. As the world grapples with the dual challenges of meeting growing energy demand and combating climate change, the harnessing of wind energy has taken center stage in the transition towards a more sustainable and environmentally responsible future [15]. Wind energy's importance lies in its ability to provide a reliable and abundant source of power without the associated greenhouse gas emissions and environmental impacts of fossil fuels. Wind turbines, strategically positioned in regions with consistent wind patterns, capture the kinetic energy of the moving air and convert it into electricity. This

---

\*Chongqing Qingdian New Energy Development Co.Ltd., Chongqing, 401121, China ([yuchenres23@outlook.co](mailto:yuchenres23@outlook.co))

process not only reduces our reliance on finite and polluting fossil fuels but also contributes to significant reductions in carbon emissions, making wind energy a critical player in mitigating climate change [14].

Moreover, the scalability of wind energy systems allows for their deployment across various scales, from small residential turbines to massive wind farms, making them adaptable to diverse energy needs. The economic benefits are also noteworthy, as the wind energy sector has spurred innovation, created jobs, and driven investment in both developed and emerging economies.

In this era of environmental consciousness and sustainability, renewable wind energy serves as a beacon of hope, offering a path towards a cleaner, more resilient energy future. This introduction sets the stage for a deeper exploration of the technologies, challenges, and opportunities associated with this essential component of our renewable energy landscape.

Solar technology spans a diverse range of applications, from rooftop solar panels on residential homes to vast solar farms that generate utility-scale electricity. Its scalability, adaptability, and eco-friendly nature make it an integral component of the global transition towards a cleaner, more sustainable energy landscape. Solar energy, often heralded as the cornerstone of sustainable power, is an increasingly prominent player in the global energy landscape. At its core, solar energy is harnessed from the most abundant and inexhaustible source known to humanity—the sun. The concept is simple yet revolutionary: by capturing the sun’s radiant energy and converting it into electricity or thermal energy, solar power offers a clean, renewable, and widely accessible solution to our growing energy needs.

The importance of solar energy cannot be overstated. In an era marked by environmental concerns, energy security, and the imperative to reduce greenhouse gas emissions, solar power emerges as a beacon of hope. By tapping into this formidable energy source, we can significantly reduce our dependence on fossil fuels, mitigate climate change, and ensure a more sustainable and resilient energy future.

The research article is organized with section 2 of objectives, section 3 of methodologies, section 4 of results and section 5 of conclusions.

## 2. Objectives.

1. To describe the impact of the scalable solution in relation to the solar and winds of the distributed generation of a parallel algorithm in the development of the smart grid environment
2. To evaluate the various or numerous methods that are used or are involved in the scalable solution or the scalability in connection with the distributed generation of solar and wind in the development of the nations
3. To explain the various advantages and benefits that are used in the solution of the scalability of the wind and solar distributed generation for a developed system of smart grid.
4. To identify the various challenges which are faced by scalable solutions in the improvement of the environment by the smart grid environment.

**3. Methodology.** The extensive explanation of the informative data as well as their analytical contemplation has been a very important part of the composition or the study of the literature. The gathering and assembling of the informative data have been done through secondary sources [22]. The gathering and collecting of the information have been verified and demonstrated in terms of their statistical and numerical data. It has been instrumental in providing better and more useful learning about the scalable solutions of winds and solar energy distributed generation for the development of the smart grid environment across global medical platforms [6]. The explanation of the scalability and smart grid, its impact, methods, and advantages that are used in improving the environment have been evaluated and explained in the study [1]. The use of qualitative data has also evolved as a valuable factor in building a strong connection between the given and elaborated information and data in the following sections.

**3.1. Impacts of the Scalable Solutions for winds and solar energy.** The term smart grid technology mainly enables the adequate and effective distribution and management of various renewable energy sources that include wind, solar, and hydrogen. The technologies of the smart grid are very useful and crucial as it connects and combines a variety of numerous distributed energy resources investment to the grid power [16]. The usage of Internet of Things (IoT) is made for the collection and assembling of informative data and is later utilized for carrying out various activities.

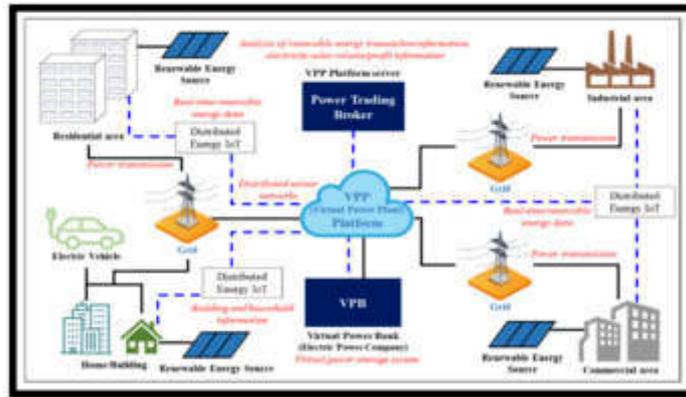


Fig. 3.1: Schematic diagram of the total energy trading system

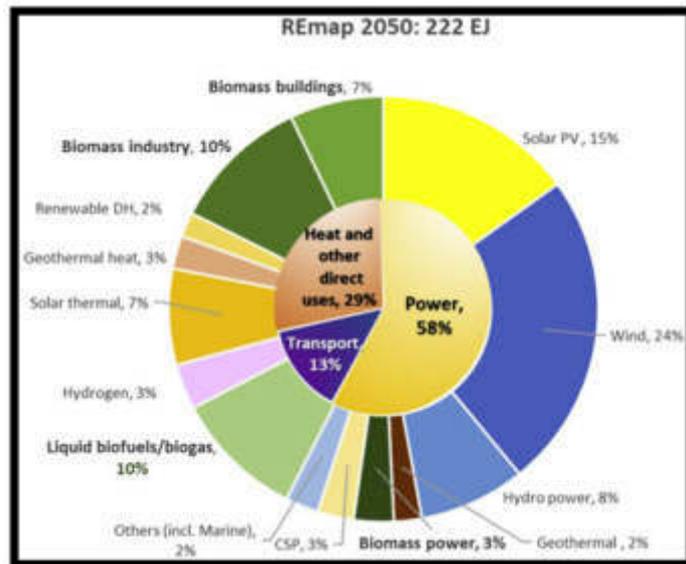


Fig. 3.2: The usage of the smart grid in renewable sources

The figure 3.1 represents the process of the smart grid system in connection with renewable sources which includes wind and solar energy. The various utilities are able to detect or catch and resolve or solve all the problems and issues quickly through continuous usages of the self-assessment [22]. These utilities no longer depend upon the consumer in order to report the outages; this process of the self-healing capacity is a very important and crucial element of the smart grid. The relationship or connections between renewable energy like solar and wind and smart grid revolve around the collecting or gathering of data [5]. This means that the farms of the winds use the mechanical gears that are vital to provide support or link to the numerous sensors. Later on this, each sensor is able to record the present environmental and climatic conditions [19]. It provides a quick response through the grid in order to give alert or warning of any problems, this enhances both the quality and the safety and services. Research has found that with the usage of the smart grid, the safety in the environment has increased by 70-80% in the environment.

The above figure 3.2 is the representation of the uses of the smart grid in renewable energy industries, which

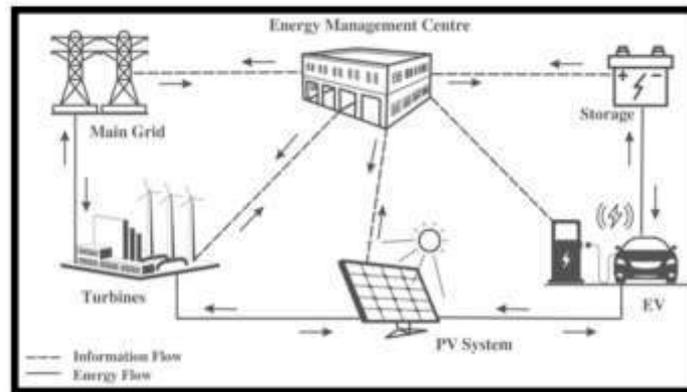


Fig. 3.3: The center of energy management

include the solar thermal industry, biomass industries, and thermal power plants [7]. In comparison with fuel-based energy power stations, renewable sources of energy also required more transformed and advanced management and distribution of power, which includes the capability for production and balancing [17]. All advancements and developed technologies can only be accomplice by the smart grid.

**3.2. Advantages of the scalable solution in relation to the solar and wind-distributed generation.** The smart grid is defined or can be explained as the improved and developed generation delivery of energy which is usually based on services that are quick and fulfill all the time demands [18]. It is mainly about the informed and influential consumption or combination that is provided by the efficiency of energy in order to reduce the environmental impacts and frequent climatic change [8]. The company named RCI is the only company that can offer or provide the complete industrial standard scalable solution for the establishment of the smart grid in the environment in connection with renewable sources of energy.

The scalable solution of the solar and wind distribution is the source of renewable energy and power. The vast distribution of electricity can be done without the burning of any kind of fuel. The resistance to the burning of fuel also reduces the rate of pollution in the environment. The US is the largest source of the production of renewable energy [13]. One of the main benefits of the use of solar panels is that the price of the solar panels is affordable. The energy and the power of the use of the solar system and scalable solution in the wind distribution generation is the high voltage of the light. The variety in the climate can be estimated properly with the help of the use of the scalable solution.

In the figure 3.3, it is a representation of the center of the management of the energy that and its process and how its transmission from the renewable energy to the various home and vehicles. The company RCI's solution mainly focused on competence in Telecommunication, IT, Advancement of Metering infrastructure, and the management of the data systems for energy which is based on the IP for the management of renewable sources of energy with the storage systems. This main aim is to create a micro-based system that is on the concept or idea of the virtual power plant for the virtual power capacity [8]. The smart grid's main objective or strategy is mostly based on distributed generation, the management of the demand and supplies, and approaches for industrial and residential, and commercial consumers. The RCI technology that is based on the introduction of the smart grid is mainly based on the IEEE smart grid conceptual model [18].

**3.3. Methods involved in the scalable solutions in connection with the solar and wind-distributed generations.** In the present times there have been seen increased demands for the integration of the distributed wind and solar resources of energy into the existing power grids that are electrical [11]. The uncertain nature of the energy resources that are renewable is the main reason for the operation of the networks that faced new issues or challenges in maintaining the balance between the generation and the load [17]. The one of the most useful methods that are used in solar technologies is the passive solar design, heating of the solar

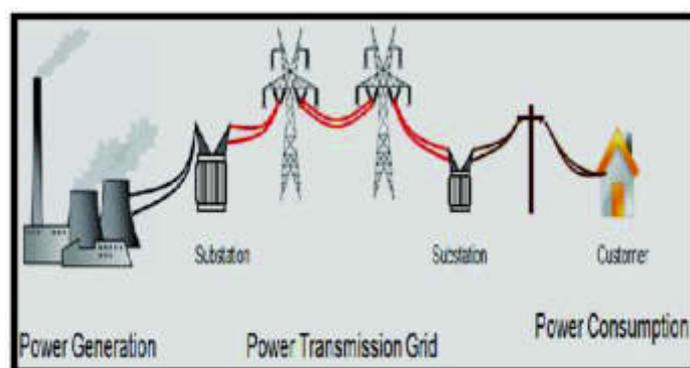


Fig. 3.4: Method involved in Grid System

water, heating and cooling of the space, and the effect of the photovoltaics.

The process of the improvement of the condition and quality of the present environment depends on the basis of the system of the solar hybrid system. The use of the photovoltaic panel of the solar system increases the capabilities and range of the batteries, charger controller, and cable connection [15]. With the help of the distribution of energy and generations, generations of a variety of technologies can happen. The best method for solar distribution is the use of a panel of solar photovoltaics. This is the simple and effortless that creates solar energy and the generation of renewable energy.

The above figure 3.4 show the methods and the techniques that are used in the grid system for the distribution of the energy. The establishment or the introduction of intelligent and smart distributed energy plants of resources are used in order to fulfill the new requirements and demands [22]. This intelligent distributed energy is used to provide the power plants which are virtually which means the management of the demand and the generation which is flexible.

**4. Main challenges faced by scalable solutions in the improvement of the environment.** The smart grid technology includes all the improved and developed real-world and digital infrastructures that provide, monitor, explain, and manage the energy or the electricity [5]. These supplies or the management of the energy are supplied from a variety of the sources in order to meet their demand for the power that is required in the homes and the business from the suppliers who supply electricity/energy [8]. This advanced and developed grid technology can supply efficiently and coordinates the demand and supply of the energy that leads to the usage of the management of the data and its capabilities. Moreover, through by the usage of advanced grid technologies, there is also a reduction in the operational costs and environmental influence in order to maximize the overall system of the stability, resilience, flexibility, and reliability.

The above figure 4.1 is the graphical representation that explains the digitalization or the advancement of the grid technologies and the initiatives taken to improve the environment by using the renewable sources like solar and wind energy [9]. The main challenges or the issues which are faced by the scalable solution for the expansion of the grid technologies in terms of renewable sources are as follows,

Research shows that the economic influence of the pandemic of the COVID-19 put efforts on the zero transmission of the harmful gases by 2050 [20]. In order to put this in effect the funding is necessary and due to the pandemic, the country has still not recovered from its losses. Restore of the national and international transport is the one of the most effective and beneficial for the recovery of the economic position of a country. The improvement of the global economy can be increased by the improvement of the global economy of each of country. The application of the SG in the various kinds of business of the countries helps to increase the rate of production and profit of the business. Improvement of the SCM is a helpful process for the recovery of the global economy. The global economy can be improved by the improvement of the connection and communication between the countries.

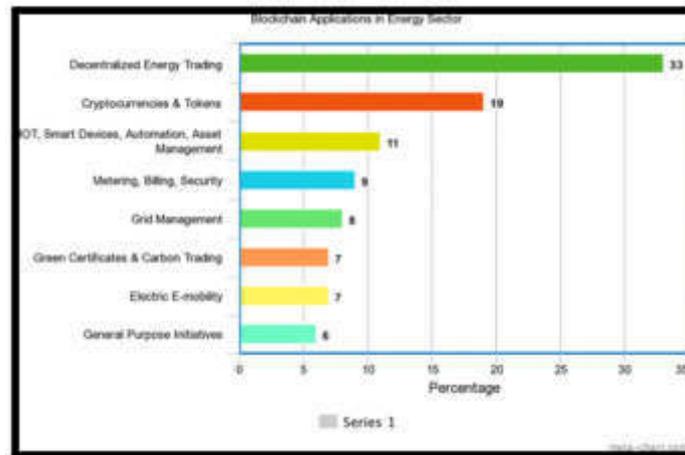


Fig. 4.1: Digitalization of the grid technologies

There are many countries that have established the farms that are related to the winds and the production of the solar power, and the zero transmission of the harmful energy generation into the grid of electricity [4]. Traditional energy like fossil fuels is still required in order to provide electrical energy to the grid power. When these countries stop the production and the usage of these fuels, this creates a scarcity of the available resources and fluctuation in the availability of the energy and the prices [2].

The fluctuation of the price and the values of the products is the reason for the decrement in the rate of import and export tax of the products. The increment of the growth of the population in the world is the reason for the enhancement of the demand for the products. The increment of the demand for the products is caused by the fluctuation of the rate and price of the products. The four main factors for the fluctuation of the price of the products are physical capital, human resources, natural resources, and the misuse of technology.

**4.1. Various Domains of the Smart Grid in Connection with renewable energy.** The National Institute of Standards and Technology (NIST) in the conceptual model of the smart grid provides or delivers a higher level of frameworks for the system of a smart grid that explains the seven vital and crucial domains, that includes, the transmission, the distribution of energy, the bulk and the advanced generations, services and market providers, customers, and operations [1]. The SG in the renewable energy covers around seven kinds of domain and the domain are related to operation, marketing, providing of the service, transmission of the energy, customer service and distribution. The operation of the interconnected domain is the factor of the capability of the operation of the various domain. The main aim of the SG technology is the upgradation of the conceptual model that remain the same kinds of the edition. The responsibilities factors of the SG help to improve the connection of the renewable energy. The improvement of the function and the electrical grid system depends on the application of the AG in order to the renewable energy.

The domain of the smart grid technology played a role for the increment of the service and communication among the stakeholders. The capacity of the decision making and the performing of the skill can be able to achieve the goal and aim of the customers. Each and every domain in itself is a combination of the all the crucial features and elements of the smart grid connecting each other with the two-way electricity, energy, and communications path [23]. These connections or the relation are basically based on the intelligent, dynamic, and future electricity grid. The performance of the service by the business and the companies helps to distribute the energy resources and decrease the load on the customers. In the each and every service sector, the proper implementation of the process and procedure is the most effective thing for the increment of the rate of production.

The above figure 4.2 represents the most important and crucial the seven domains which are involve in the

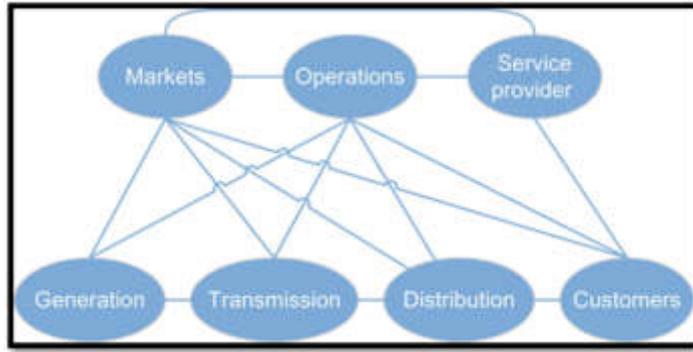


Fig. 4.2: The Seven Domains of Smart Grid

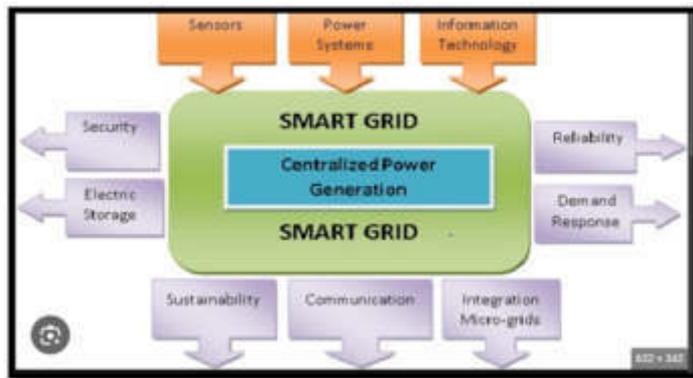


Fig. 4.3: The benefits of the Smart Grid System

grid system. These seven and the most important domains are how all the communications and the flow of the electricity and the energy combines or connect each domain [12]. These are all domains that influence the flow of electricity and are vital and as well as connected to each other [21].

**4.2. Results.** According to the study, it has been shown that the establishment of the smart grid market is expected to reach 80 billion by 2024. The growth and the development in the inclination towards digitalization or enhancement of the grid energy supply system are nearly owned 50% of the market growth and supply 60-50% of the energy or electricity that is not harmful to the environment. Most of the industries face issues and challenges due to the power cut and declining resilience from the traditional transmission and production infrastructure. In order to modernize the activities of the grid many countries have been preparing to deploy the infrastructure of the smart grid.

The deployment or the destruction of the technologies which are related to the smart grid offers or provides many advantages for the combination of the renewable resource of energy that includes solar and wind energy. Therefore, this grid technology facilitates and benefits the increased demand for the security of energy. In addition, there are many favorable schemes and the initiative has been taken to transform the energy generation and its infrastructure that will develop the smart grid by making it a global market trend by 2025.

As a result, from the above study, it is clear that the foundation of SG technology in the business sector provides the evolution of the power grids that exist in the present day. The increment of the security and privacy and the reliability of the grids that are traditional are mostly impracticable. The process of the solve of the issues and problems of the SG puts pressure on the huge amount of energy consumption and the backbone

of the bandwidth of the business. The monitoring of the real-time and the controlling of useful applications are two main important factors of the use of the SG technologies. The addition of the EC to the SG helps to solve the issues and problems and improve the efficiency of energy consumption. From the research, it is clear that the installation of IoT-based technologies in the solar distribution techniques helps to decrease the issues and problems and improve the application of the parallel algorithms in the solar distribution.

**5. Conclusion.** In conclusion, the results of this study shed light on the immense potential and importance of the smart grid system in the evolving energy landscape. The research highlights that the smart grid market is on a trajectory to reach an impressive 80 billion by 2024, signifying the significant growth and development in the transition towards a more digital and sustainable energy supply system. Furthermore, the study underscores the pivotal role of smart grid technology in addressing the limitations of traditional power grids. It offers heightened security, improved privacy, and enhanced reliability, all of which are increasingly impracticable with conventional grids. However, it is essential to acknowledge the energy consumption and bandwidth demands that come with solving these issues. Real-time monitoring and control are crucial facets of smart grid technology, and the incorporation of Energy Cloud (EC) solutions further enhances energy efficiency. One of the key takeaways from this study is the critical role of smart grid technology in promoting the integration of renewable energy sources, particularly solar and wind power. This integration not only bolsters the global shift towards greener energy but also caters to the increasing demand for energy security. The adoption of favorable schemes and initiatives aimed at transforming energy generation and infrastructure positions the smart grid as a prominent global market trend by 2025.

#### REFERENCES

- [1] I. AHMED, M. REHAN, A. BASIT, S. H. MALIK, K.-S. HONG, ET AL., *Multi-area economic emission dispatch for large-scale multi-fueled power plants contemplating inter-connected grid tie-lines power flow limitations*, Energy, 261 (2022), p. 125178.
- [2] S. J. ALSUNAIDI AND F. A. KHAN, *Blockchain-based distributed renewable energy management framework*, IEEE Access, 10 (2022), pp. 81888–81898.
- [3] N. ALTIN AND S. E. EYIMAYA, *Artificial intelligence applications for energy management in microgrid*, in 2023 11th International Conference on Smart Grid (icSmartGrid), IEEE, 2023, pp. 1–6.
- [4] X. CHEN, G. QU, Y. TANG, S. LOW, AND N. LI, *Reinforcement learning for selective key applications in power systems: Recent advances and future challenges*, IEEE Transactions on Smart Grid, 13 (2022), pp. 2935–2958.
- [5] H. M. HUSSAIN, A. NARAYANAN, P. H. NARDELLI, AND Y. YANG, *What is energy internet? concepts, technologies, and future directions*, Ieee Access, 8 (2020), pp. 183127–183145.
- [6] A. JOSEPH AND P. BALACHANDRA, *Smart grid to energy internet: A systematic review of transitioning electricity systems*, IEEE Access, 8 (2020), pp. 215787–215805.
- [7] R. KHALID, N. JAVAID, A. ALMOGREN, M. U. JAVED, S. JAVAID, AND M. ZUAIR, *A blockchain-based load balancing in decentralized hybrid p2p energy trading market in smart grid*, Ieee Access, 8 (2020), pp. 47047–47062.
- [8] R. L. KINI, D. RAKER, R. MARTIN-HAYDEN, R. G. LUTES, S. KATIPAMULA, R. ELLINGSON, M. J. HEBEN, AND R. KHANNA, *Control-centric living laboratory for management of distributed energy resources*, IEEE Open Access Journal of Power and Energy, 10 (2022), pp. 48–60.
- [9] X. KOU, F. LI, J. DONG, M. STARKE, J. MUNK, Y. XUE, M. OLAMA, AND H. ZANDI, *A scalable and distributed algorithm for managing residential demand response programs using alternating direction method of multipliers (admm)*, IEEE Transactions on Smart Grid, 11 (2020), pp. 4871–4882.
- [10] Y. LI AND J. YAN, *Cybersecurity of smart inverters in the smart grid: A survey*, IEEE Transactions on Power Electronics, (2022).
- [11] M. MASSAOUDI, H. ABU-RUB, S. S. REFAAT, I. CHIHI, AND F. S. OUESLATI, *Deep learning in smart grid technology: A review of recent advancements and future prospects*, IEEE Access, 9 (2021), pp. 54558–54578.
- [12] M. MELIANI, A. E. BARKANY, I. E. ABBASSI, A. M. DARCHERIF, AND M. MAHMOUDI, *Energy management in the smart grid: State-of-the-art and future trends*, International Journal of Engineering Business Management, 13 (2021), p. 18479790211032920.
- [13] Q. N. MINH, V.-H. NGUYEN, V. K. QUY, L. A. NGOC, A. CHEHRI, AND G. JEON, *Edge computing for iot-enabled smart grid: The future of energy*, Energies, 15 (2022), p. 6140.
- [14] M. B. MOLLAH, J. ZHAO, D. NIYATO, K.-Y. LAM, X. ZHANG, A. M. GHAS, L. H. KOH, AND L. YANG, *Blockchain for future smart grid: A comprehensive survey*, IEEE Internet of Things Journal, 8 (2020), pp. 18–43.
- [15] I. MURZAKHANOV, S. GUPTA, S. CHATZIVASILEIADIS, AND V. KEKATOS, *Optimal design of volt/var control rules for inverter-interfaced distributed energy resources*, IEEE Transactions on Smart Grid, (2023).
- [16] N. I. NIMALSIRI, C. P. MEDIWATHTHE, E. L. RATNAM, M. SHAW, D. B. SMITH, AND S. K. HALGAMUGE, *A survey of algorithms for distributed charging control of electric vehicles in smart grid*, IEEE Transactions on Intelligent Transportation Systems, 21 (2019), pp. 4497–4515.

- [17] A. OTHMAN, G. KADDOUM, J. V. EVANGELISTA, M. AU, AND B. L. AGBA, *Digital twinning in smart grid networks: Interplay, resource allocation and use cases*, IEEE Communications Magazine, (2023).
- [18] C. PARTHASARATHY, H. HAFEZI, H. LAAKSONEN, AND K. KAUHANIEMI, *Modelling and simulation of hybrid pv & bes systems as flexible resources in smartgrids—sundom smart grid case*, in 2019 IEEE Milan PowerTech, IEEE, 2019, pp. 1–6.
- [19] I. PETRI, M. BARATI, Y. REZGUI, AND O. F. RANA, *Blockchain for energy sharing and trading in distributed prosumer communities*, Computers in Industry, 123 (2020), p. 103282.
- [20] C. QIN, A. K. SRIVASTAVA, AND K. L. DAVIES, *Unbundling smart meter services through spatiotemporal decomposition agents in der-rich environment*, IEEE Transactions on Industrial Informatics, 18 (2021), pp. 666–676.
- [21] D. SYED, A. ZAINAB, A. GHAYEB, S. S. REFAAT, H. ABU-RUB, AND O. BOUHALI, *Smart grid big data analytics: Survey of technologies, techniques, and applications*, IEEE Access, 9 (2020), pp. 59564–59585.
- [22] W. WANG, X. FANG, H. CUI, F. LI, Y. LIU, AND T. J. OVERBYE, *Transmission-and-distribution dynamic co-simulation framework for distributed energy resource frequency response*, IEEE Transactions on Smart Grid, 13 (2021), pp. 482–495.
- [23] R. ZAFAR, A. E. NEZHAD, A. AHMADI, T. ERFANI, AND R. ERFANI, *Trading off environmental and economic scheduling of a renewable energy based microgrid under uncertainties*, IEEE Access, 11 (2022), pp. 459–475.

*Edited by:* Sathishkumar V E

*Special issue on:* Scalability and Sustainability in Distributed Sensor Networks

*Received:* Aug 25, 2023

*Accepted:* Oct 30, 2023



## PHYSICAL EDUCATION TEACHING QUALITY EVALUATION METHOD USING MOBILE EDGE COMPUTING IN THE ONLINE AND OFFLINE ENVIRONMENT

HUIMIN ZHANG\*

**Abstract.** The development of different technology has impacted the different stages of human life. Additionally, with the implication of computing technology sports and physical education can be improved. The use of mobile edge technology aids in gathering precise data related to physical education. Therefore, a specific improvement is possible. This following analysis has looked into the factor of mobile edge computing that aids in the evolution of physical education teaching. Moreover, the study has focused on developing an appropriate path for the integration of MEC into PE education. By utilising Mobile Edge Computing (MEC) technologies in both online and off-line learning contexts, this study offers a thorough way for assessing the quality of sports teaching. The suggested approach integrates multiple characteristics and performance measures to evaluate the efficacy of physical education instruction and participation, taking into consideration the changing educational landscape.

**Key words:** Mobile edge technology, Physical education, issues in implementing MEC for PE

**1. Introduction.** The development of physical education (PE) teaching has a vast number of factors. For instance, there are physical as well as mental benefits of quality PE training [4]. Therefore, the following empirical analysis has looked into the factor of physical education teaching quality evaluation method using mobile edge computing in the online and offline environment [19]. Furthermore, through a systematic analysis all the possible factors of the same are discussed and the results are developed based on the analysis of secondary information. Therefore, a detailed analysis is presented in the following study.

Physical education plays a fundamental role in promoting physical health and well-being. It helps individuals develop fitness, strength, and endurance, reducing the risk of health issues such as obesity, heart disease, and diabetes [3]. It provides a platform for students to acquire and refine a wide range of motor skills, from basic movements to more complex sports-specific techniques. These skills contribute to overall physical competence. Physical education instills the importance of regular physical activity, encouraging students to adopt and maintain a healthy lifestyle throughout their lives. It fosters habits of exercise and fitness. Regular physical activity is associated with improved mental health [5]. Physical education can reduce stress, anxiety, and depression while enhancing mood and overall psychological well-being. Physical education fosters social interaction and teamwork. Participation in sports and group activities helps students develop essential social skills, such as communication, cooperation, and sportsmanship. It teaches discipline, responsibility, and time management. Students learn the importance of punctuality, preparation, and adhering to rules, which are valuable life skills [22].

The research offers a comprehensive approach to assess the quality of sports teaching. By integrating multiple characteristics and performance measures, it provides a robust framework for evaluating the effectiveness of physical education instruction and participation. This contribution is valuable for ensuring that the evolving educational landscape is well-catered to.

This research contributes as follows,

1. The study recognizes the impact of technological advancements on various aspects of human life and extends this impact to sports and physical education.
2. The research identifies the utility of MEC technology in gathering precise data related to physical education. This contribution underscores the significance of data-driven insights in enhancing the teaching and learning experiences in sports and physical education.

---

\*School of Public Teaching and Practice Wuhan Technical College of Communications, Wuhan, 430065, China (huiminzhang1@outlook.com)

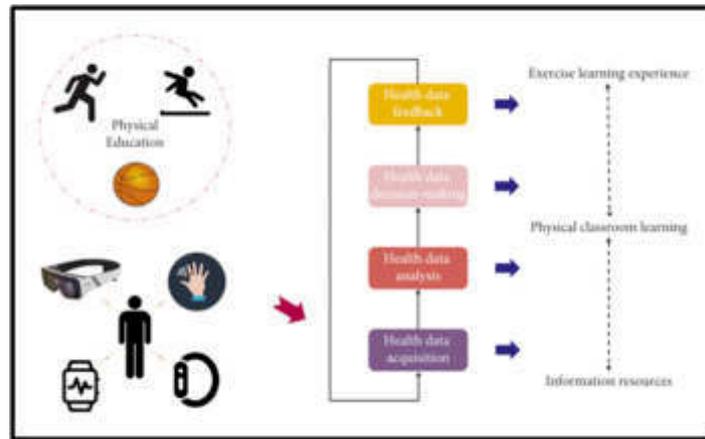


Fig. 1.1: Framework for technology-based PE education

**2. Objectives.** In order to develop the analysis in a coherent manner following objectives were followed:

1. To discuss the possible factors related to the development of Physical education using mobile edge computing
2. To look into different technology that aids in the development process of Physical education teaching
3. To analyse possible issues in the implication of mobile edge technology in the development process of Physical education teaching
4. To discuss the possible solution for the identified problems

**3. Methodology.** The methodology of the following analysis deals with different elements and strata of the empirical analysis. Moreover, all the techniques used in the process are depended on the methodology of the study [5]. For the comprehension of factors and issues related to physical education, all the possible factors are analysed through the collection of secondary information. Secondary data aid in the development of research which is ethically sound and delta oriented [22]. Moreover, the process of developing a reliable result is supported by the use of qualitative analysis. Therefore, for the holistic development of the study, a secondary qualitative method was employed.

**3.1. Elements related to the evolution of PE teaching using mobile edge computing .** At the time of analysing past literature related to computing technology for PE education, it was observed that there are certain factors that directly or indirectly impact the implication. Following are some of the factors that are related to the implication of computing technology for the evolution of PE teaching:

*Collection and processing of real-time information.* MEC enables real-time data collection and analysis of the same [23]. Therefore, enabling the gathering and analysis of movement data from students during workouts aid to contemplate the issues in the training.

Moreover, such information can be utilized to deliver individualized feedback, monitor development, and instantly change the complexity of tasks [14] Additionally, classified training programs can be formulated with the help of such specified information.

*Interactive session for teaching PE.* MEC is a whole system that has interactive elements such as augmented reality (AR) and virtual reality (VR) experiences. In addition, with the implication of MCA for PE education having a realistic experience is possible [16]. Hence, by generating realistic virtual settings in which students may partake in a variety of physical activities, these technologies might increase students' motivation and engagement.

The effect of the interactive session between the PE teacher and the students is undeniable for the physical and mental growth of the students. With the help of proper integration and communication, the issues and problems of the students can be determined properly. The issues for the lack of understanding of the students can be find by integrating with the students. The use of the mobile computing programming and online sources

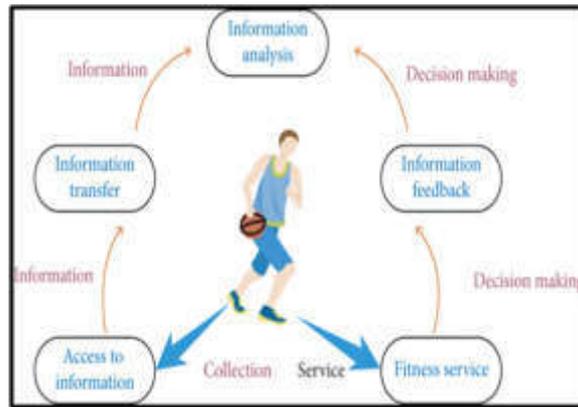


Fig. 3.1: Process of data collection and decision making for PE

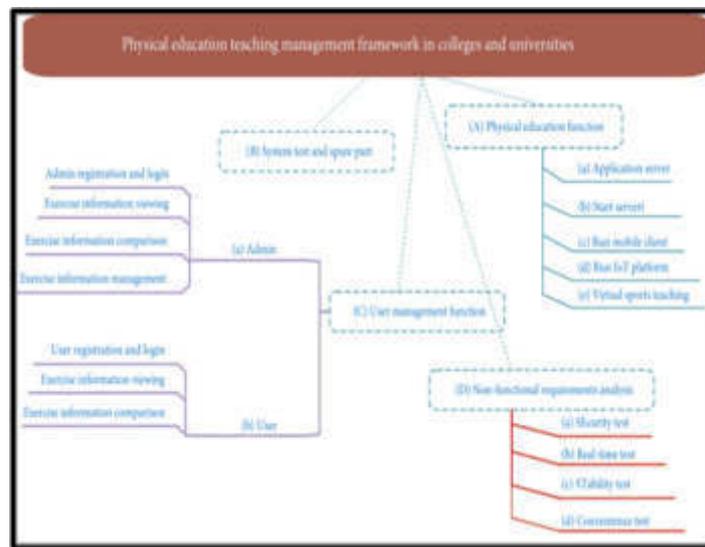


Fig. 3.2: Framework for implementing MCE technology in universities and colleges

help to determine the personal issues of the students [25]. Giving the examples using the virtual tools attract the students to see the examples and sometimes the examples help the students to come out from the mental disorder. The use of the online system in the educational department increases the level of attraction of the students and this effects on the willingness of the students. Therefore, the application of the online and offline strategies are helpful for the growth and development of the students.

*Appropriate resource allocation.* During the analysis of the past literature, it was noted that process resource allocation is a major factor that influences the MEC implication for PE. Moreover, it was noted that appropriate

*Teachers' knowledge and understanding.* It was noted that the technological understanding of teachers or coaches is essential for the effective implication of MEC technology [20]. Moreover, making decisions based on the information collected from the MEC technology is an important part of the evolution of teaching. Hence, training of the teachers in relation to technology is essential.

*Data security.* During the past data analysis it was noted that the data is an essential aspect of the evolution and implication of MEC technology [2]. Moreover, the data breach can damage the implication of the evolution of teaching. Therefore, securing the information is essential for such instances.

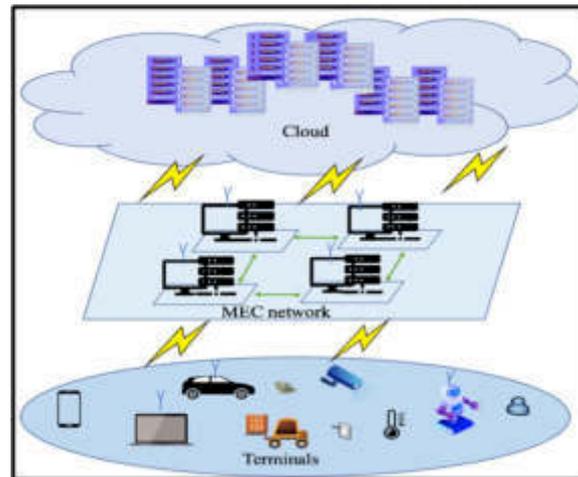


Fig. 3.3: Architecture of MEC-based IoT technology

**3.2. The real-time implication of MEC in physical education.** The implementation of the MEC in the section of the physical education helps the students for their better improvement and grow. The development of the physical growth of the students also includes the growth of the physical, mental and emotional. One of the most vital and important factors of the physical department is the maintenance of the discipline of the students. Discipline is the most vital and important factor for the growth and development in the life of every individual [19]. The practice of uncontrolled behavior becomes one of the effected on the future of the students. Therefore, the controlling of uncontrolled behavior can be done by the application of the physical training and discipline of the students.

The concept of physical movements is affected by the exploration of the awareness of the body and increases the relationship among the different parts of the body. The use of the mobile device and the internet system in the physical education system helps to increase awareness among the students [16]. The increment of awareness about physical movement is a helpful and effective process for the future growth of the students.

The figure 3.3 of the empirical analysis shows a detailed working process of MEC technology. All of the components of the technology can be seen in action in the above illustration. MEC technology works with integration with IoT devices in order to work in an appropriate manner [6]. Moreover, the network can be divided into three layers that are cloud storage, MEC networks, and terminals. The terminal works as a mediator station and provides actionable data in the terminals where actions are taken [19]. In addition, the server of an MCE technology acts as a storage for the data. Moreover, the server stores information thus accurate historical data aids in the comparison of information.

Moreover, the process of simulation data indicated the accuracy of the system in order to improve training for PE education [13]. Therefore, it can be contemplated that with the implication of MEC technology improvement in the development process is possible. Therefore, developing an accurate system in online as well as offline mode is possible. Moreover, the process aids to determine the vulnerability of an athlete and points out the gaps for the player.

**3.3. Issues related to the implication of MEC for the evolution of PE training.** Through the analysis of past information, it was noted that there are certain issues related to the implication of MEC for PE education. Such issues hinder the development of PE training and resource wastage is common in such situations [12]. Following is a coherent discussion of the issues related to the implication of MEC for PE education

*Financial resources and accessibility.* One of the major issues that can hinder the implication of MEC for PE education is found to be financial resources [25]. The development of MEC infrastructure requires different elements which are expensive and expensive. Lack of funding makes it difficult for institutions or schools to

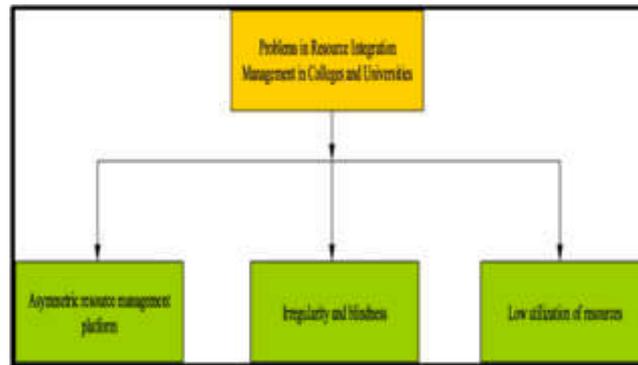


Fig. 3.4: Problems of resource allocation for MEC-based PE education

purchase the required tools and technology, which might result in an uneven distribution of MEC-enhanced PE activities [1].

*Lack of adequate knowledge of human resources.* To utilize MEC-based machines and platforms successfully, both teachers and students need to get training [11]. Moreover, a basic meaning for handling systems and comprehending information is essential for human resources. Thus, the lack of training can hamper the appropriate use of such technology [18]. In addition, the influence on PE instruction can be limited by a lack of technical skills.

*Infrastructure development.* The network infrastructure that MEC depends required to be strong and dependable [7]. Network coverage issues in some areas might result in differences in who has access to MEC-based PE instruction. For a flawless experience, it is crucial to have good connectivity [21]. Hence, it was noted that infrastructure is the basic issue that needs to be improved for an effective evolution of the PE system.

The development of the infrastructure of the education institute and the schools helps the students in their future growth. Developing the infrastructure of the educational institute helps the students to understand the usefulness of the learning capabilities and activities. The recruitment of the skilled and experience PE instructors in the educational department helps the students to grow the physical skilled. For the physical improvement of the students, the implementation of the modern and innovative tools is very much important for the students of the physical department [22]. Thus, the value and effect of the improvement of the infrastructure of the educational sectors is helpful for the growth and development of the students.

*Real-time information.* One of the major benefits of MEC technology is that it provides a real-time perspective for PE education [6]. In addition, real-time information aids to analyse gaps and adapt to the situation based on real-time data. Therefore, it can be contemplated that real-time information provided by the MEC technology aid in the evolution of PE education teaching.

**4. Results.** In order to understand the impact of MEC technological enhancement on the evolution of PE teaching a secondary analysis was conducted. Moreover, through the systematic analysis of past information, it was contemplated that there is a high chance of improvement in PE education [24]. It was noticed that there are certain elements in the possess of implementing MEC technology for PE education. For instance, the collection and processing of real-time information is a major factor in the evolution. Similarly, it was noted that training and understanding of the teachers influence the decision-making process for PE education [15].

As a result, it can be said that the use of the HBCR to assist the wearable sensors has a positive effect on the increment of the quality of teaching. Application of the aerobic training is the most beneficial and helpful for the educational growth of the students of physical education. The comparison between the CRF and CBCR has no significance difference and the limitation of the symptom helps to test the observe group [9]. From the whole study, this can be said that the outcomes of the use if CPET for the students of the physical education made a great difference between the other students. The unusual participation of the students in the aerobic training decreases the specification and intensity of the students. For analysing the information related to the

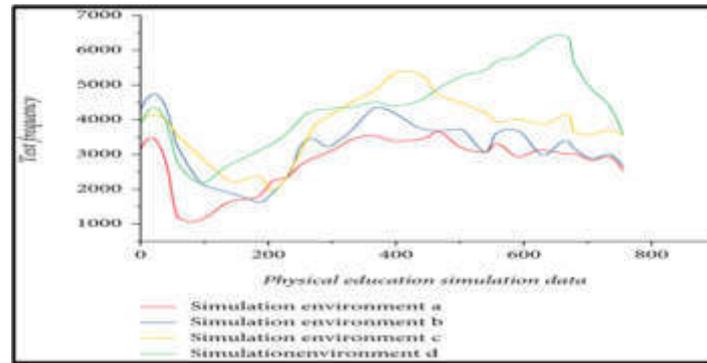


Fig. 4.1: Simulation data of different environments

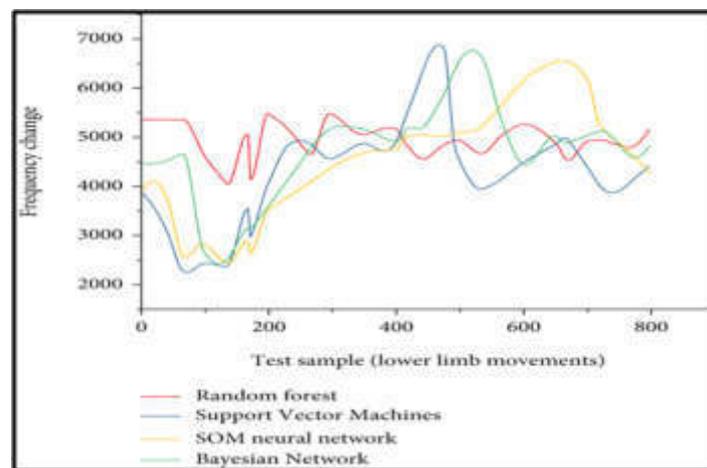


Fig. 4.2: Simulation data of lower limb

implication of MEC for PE a real-time analysis was conducted. For the study, it was noted that their different environments have different implications for the MEC system for PE education [17]. At the moment, migration learning is required to make edge computing end servers cooperate [9].

The application of the mobile edge computing and the online system increase the capacity of the understanding of the students. For the students of the physical education, the proper demonstration of the physical activity and the outcomes of the activity helps to increase the attention of the students. With the help of the use of the mobile edge and the internet process, the education authority can be able to demonstrate the effect and impact of the practice of the activity of the physical department of education [1]. The use of the offline training physical students is the most effective to improve the quality of the teaching learning process. The self-learning is the most effective and valuable for the growth and development of the students who studied the physical education.

The figure 4.1 depicts the simulation information related to physical education. For the study, 4 different environments subjugated the overall test. Moreover, the implication of simulation data shows that the majority of the system's functionalities can satisfy the system requirements [7].

The figure 4.2 is related to the low limb movement frequency changes. It can be seen that the development of the system's fundamental purpose can handle the administration of physical education [7]. Additionally, it can be contemplated that a lot of sports event publicity can be immediately published to the system with the aid of such a method. Therefore, the value of the sports education management system is progressively represented

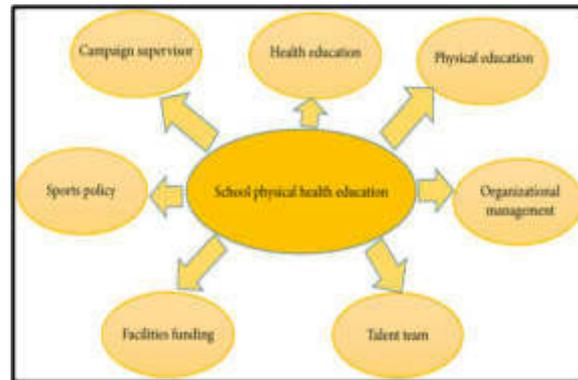


Fig. 4.3: Different dimensions of PE in schools

in terms of sports event registration, score entry, and associated sports event publicity [10]. Moreover, with the implication of complex algorithms, a better perspective can be achieved for the test.

The simulation of the use of the offline environment in the physical education institute effect on the future development of the children. Sometimes the teachers and mentors of the educational institute takes the students in a educational excursion which boost the experience of the students and improve the level of understanding of the students. The application of the super vector machine in the educational department is the one of the most helpful and effective process and strategies for the betterment of the students [17]. The neural network is the most beneficial for the implementation of the innovative technology in the physical department of the physical educational institute. Thus, in the present days, the use of the online and offline computing practice helps to increase the skilled of the students.

The figure 4.3 provides a detailed illustration related to PE in school. It can be seen that the impact of physical education can be traced to different aspects of school education. Furthermore, through the secondary analysis of environmental simulation, it was observed that the environment has a major impact on the implication process [8]. Moreover, the efficiency of the evolution process for the teaching of PE education has a direct relation with the environment. It was analysed that there are certain aspects that create a hindrance to the evolution process. Moreover, such hindrances have an influence on online as well as offline mediums [9]. For instance, understanding the staff member has a major influence on the implication process. Additionally, the expensive nature of the system influences the process of implication

**5. Conclusion.** Thus, the above analysis has coherently discussed information related to the use of MEC technology in order to develop the PE educational strategy. Moreover, the process has highlighted various factors related to the technology and its implication. It was established that various environments have various impacts on the implication of MEC. Furthermore, it was established that the factors such as cost and efficiency of human resources have a direct impact on the implication. Therefore, with effective use of the system it was noted that certain evolution for the PE training is possible. Additionally, both online and offline modes of training are impacted by such factors. The study emphasizes the potential positive impact of MEC technology on the evolution of PE education. It highlights the role of real-time data collection and processing as a crucial factor in this evolution. This implies that the use of MEC can enhance the quality and effectiveness of teaching in the field of physical education. The study points out the effectiveness of utilizing Heart Rate-Based Cognitive Radio (HBCR) in conjunction with wearable sensors for improving the quality of teaching. In particular, the application of aerobic training is highlighted as a valuable educational tool. Real-time analysis is identified as a critical tool for understanding the implications of MEC for PE education. The study recognizes that different environments can have varying effects on the MEC system's performance, indicating the need for adaptability and customization.

## REFERENCES

- [1] M. AAZAM, S. ZEADALLY, AND E. F. FLUSHING, *Task offloading in edge computing for machine learning-based smart healthcare*, *Computer Networks*, 191 (2021), p. 108019.
- [2] M. G. R. ALAM, M. M. HASSAN, M. Z. UDDIN, A. ALMOGREN, AND G. FORTINO, *Autonomic computation offloading in mobile edge for iot applications*, *Future Generation Computer Systems*, 90 (2019), pp. 149–157.
- [3] K. BAJAJ, B. SHARMA, AND R. SINGH, *Implementation analysis of iot-based offloading frameworks on cloud/edge computing for sensor generated big data*, *Complex & Intelligent Systems*, 8 (2022), pp. 3641–3658.
- [4] L. BAO AND P. YU, *Evaluation method of online and offline hybrid teaching quality of physical education based on mobile edge computing*, *Mobile Networks and Applications*, 26 (2021), pp. 2188–2198.
- [5] A. BOURECHAK, O. ZEDADRA, M. N. KOUAHLA, A. GUERRIERI, H. SERIDI, AND G. FORTINO, *At the confluence of artificial intelligence and edge computing in iot-based applications: A review and new perspectives*, *Sensors*, 23 (2023), p. 1639.
- [6] Y. CHEN, Z. LIU, Y. ZHANG, Y. WU, X. CHEN, AND L. ZHAO, *Deep reinforcement learning-based dynamic resource management for mobile edge computing in industrial internet of things*, *IEEE Transactions on Industrial Informatics*, 17 (2020), pp. 4925–4934.
- [7] Q. CUI, Z. GONG, W. NI, Y. HOU, X. CHEN, X. TAO, AND P. ZHANG, *Stochastic online learning for mobile edge computing: Learning from changes*, *IEEE Communications Magazine*, 57 (2019), pp. 63–69.
- [8] X. DENG, Z. SUN, D. LI, J. LUO, AND S. WAN, *User-centric computation offloading for edge computing*, *IEEE Internet of Things Journal*, 8 (2021), pp. 12559–12568.
- [9] A. DEV ET AL., *Iot and fog computing based prediction and monitoring system for stroke disease*, *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, 12 (2021), pp. 3211–3223.
- [10] B. FAN, Y. WU, Z. HE, Y. CHEN, T. Q. QUEK, AND C.-Z. XU, *Digital twin empowered mobile edge computing for intelligent vehicular lane-changing*, *IEEE Network*, 35 (2021), pp. 194–201.
- [11] S. IFTIKHAR, S. S. GILL, C. SONG, M. XU, M. S. ASLANPOUR, A. N. TOOSI, J. DU, H. WU, S. GHOSH, D. CHOWDHURY, ET AL., *Ai-based fog and edge computing: A systematic review, taxonomy and future directions*, *Internet of Things*, (2022), p. 100674.
- [12] F. LIN, Y. ZHOU, I. YOU, J. LIN, X. AN, AND X. LÜ, *Content recommendation algorithm for intelligent navigator in fog computing based iot environment*, *IEEE Access*, 7 (2019), pp. 53677–53686.
- [13] T. LONG, Y. MA, Y. XIA, X. XIAO, Q. PENG, AND J. ZHAO, *A mobility-aware and fault-tolerant service offloading method in mobile edge computing*, in *2022 IEEE International Conference on Web Services (ICWS)*, IEEE, 2022, pp. 67–72.
- [14] Y. NAN, W. LI, W. BAO, F. C. DELICATO, P. F. PIRES, Y. DOU, AND A. Y. ZOMAYA, *Adaptive energy-aware computation offloading for cloud of things systems*, *IEEE access*, 5 (2017), pp. 23947–23957.
- [15] A. SHAHIDINEJAD, F. FARAHBAKHSH, M. GHOBAEI-ARANI, M. H. MALIK, AND T. ANWAR, *Context-aware multi-user offloading in mobile edge computing: a federated learning-based approach*, *Journal of Grid Computing*, 19 (2021), pp. 1–23.
- [16] J. SHENG, J. HU, X. TENG, B. WANG, AND X. PAN, *Computation offloading strategy in mobile edge computing*, *Information*, 10 (2019), p. 191.
- [17] M. S. SOFLA, M. H. KASHANI, E. MAHDIPOUR, AND R. F. MIRZAEI, *Towards effective offloading mechanisms in fog computing*, *Multimedia Tools and Applications*, 81 (2022), p. 1997.
- [18] O. TAO, X. CHEN, Z. ZHOU, L. LI, AND X. TAN, *Adaptive user-managed service placement for mobile edge computing via contextual multi-armed bandit learning*, *IEEE Transactions on Mobile Computing*, (2021).
- [19] L. WANG AND Z. XU, *An english learning system based on mobile edge computing constructs a wireless distance teaching environment*, *Mobile Information Systems*, 2021 (2021), pp. 1–9.
- [20] T. WANG, Y. LIANG, Y. ZHANG, X. ZHENG, M. ARIF, J. WANG, AND Q. JIN, *An intelligent dynamic offloading from cloud to edge for smart iot systems with big data*, *IEEE Transactions on Network Science and Engineering*, 7 (2020), pp. 2598–2607.
- [21] Z. WANG, T. LV, AND Z. CHANG, *Computation offloading and resource allocation based on distributed deep learning and software defined mobile edge computing*, *Computer Networks*, 205 (2022), p. 108732.
- [22] X. XU, Q. LIU, Y. LUO, K. PENG, X. ZHANG, S. MENG, AND L. QI, *A computation offloading method over big data for iot-enabled cloud-edge computing*, *Future Generation Computer Systems*, 95 (2019), pp. 522–533.
- [23] A. YOUSAFZAI, I. YAQOUB, M. IMRAN, A. GANI, AND R. M. NOOR, *Process migration-based computational offloading framework for iot-supported mobile edge/cloud computing*, *IEEE internet of things journal*, 7 (2019), pp. 4171–4182.
- [24] S. K. U. ZAMAN, A. I. JEHangIRI, T. MAQSOOD, N. U. HAQ, A. I. UMAR, J. SHUJA, Z. AHMAD, I. B. DHAOU, AND M. F. ALSHAREKH, *Limpo: Lightweight mobility prediction and offloading framework using machine learning for mobile edge computing*, *Cluster Computing*, 26 (2023), pp. 99–117.
- [25] L. ZHANG, P. WEI, Y. ZHANG, N. WANG, ET AL., *Artificial intelligence and edge computing technology promote the design and optimization of flipped classroom teaching models for higher vocational, ideological, and political courses*, *Mobile Information Systems*, 2022 (2022).

*Edited by:* Sathishkumar V E

*Special issue on:* Scalability and Sustainability in Distributed Sensor Networks

*Received:* Aug 25, 2023

*Accepted:* Oct 29, 2023



## STABILITY STUDY OF GRID-CONNECTED POWER SYSTEM FOR WIND FARMS CONSIDERING POWER CONTROL

LIANGNIAN LV, ZHONG FANG, ZHIQIAN YANG, AISIKAER, LEI WU\*, AND YANCHEN YANG<sup>†</sup>

**Abstract.** Wind energy has emerged as a pivotal practice in the contemporary energy landscape, generated through grid-connected power sources aligning with the vernacular principles of systemic approaches. This study explores the surge in initiation rates, offering insights into various factors impacting sustainable electricity production. Intriguingly, this research delves into the intricacies of managing the variability and uncertainty inherent in energy demand, catalysing the integration of grid-based solutions that enhance sustainability. It probes the dynamic nature of power supply paradigms, revealing a journey of continuous enhancement by applying cutting-edge resource methodologies. Amidst the backdrop of global shifts in electricity dynamics, this study uncovers the profound implications of energy depletion and wasteful consumption practices, spotlighting a burgeoning movement towards optimising grid electricity resources on a macro scale. The intricacies and nuances of power supply challenges are comprehensively dissected, offering valuable insights. Furthermore, the study explores the pivotal role played by information technology innovators in consolidating the predictability of wind energy, augmenting its viability. It also aligns with forward-looking reviews, underscoring the actionable strategies taken. The culmination of these efforts not only enhances predictability but also unlocks a spectrum of reflective and adaptive resources in wind energy utilisation.

**Key words:** Grid frequency, variability, wind farms, AC-DC converters, modelling, power control

**1. Introduction.** The control of grid frequency is an essential power system that is used to penetrate wind turbines for maintaining the stability of power grids. By allowing individuals to access and evaluate the source code, the security level increases. The evaluation of such designs is important in terms of predicting the conditions in considering the grid frequency.

The importance and motivation of this research are multifaceted and underscore its significance within the broader scientific and technological landscape. Firstly, this study addresses the pressing global need for sustainable energy sources. In an era characterized by environmental concerns and the imperative to reduce carbon emissions, wind energy stands out as a clean and renewable solution. Understanding its dynamics and harnessing its potential is crucial for mitigating climate change and ensuring a more sustainable future. Moreover, the motivation behind this research is driven by the ever-increasing demand for electricity. As our societies become more technologically advanced and interconnected, the need for reliable and efficient energy sources becomes paramount. Wind energy has the potential to contribute significantly to our energy mix, but its variability and unpredictability pose challenges that this research aims to address. By exploring methods to enhance the predictability and reliability of wind energy, this study seeks to ensure a stable and consistent power supply, reducing our dependence on fossil fuels and their associated environmental consequences.

Furthermore, the research's importance lies in its contribution to technological innovation. As we grapple with the complex challenges of integrating intermittent renewable energy sources into our power grids, the development of modern methods and solutions becomes essential. This research explores these cutting-edge approaches, offering potential solutions that can revolutionize how we harness wind energy. By doing so, it not only addresses pressing environmental concerns but also contributes to the evolution of energy systems and grid management. It strives to make wind energy a more reliable and predictable component of our energy infrastructure, ultimately driving us toward a greener and more sustainable energy future.

### 2. Objectives.

1. To analyze the role of grid electricity in managing booth variability and uncertainty of loads
2. To determine the importance of generator modeling and wind turbine

---

\*Goldwind Science & Technology Co., Ltd, Beijing, China,100176 ([liangninlv1@outlook.com](mailto:liangninlv1@outlook.com))

<sup>†</sup>State Grid Jibei Electric Power Co., Ltd, Beijing, China, 100032

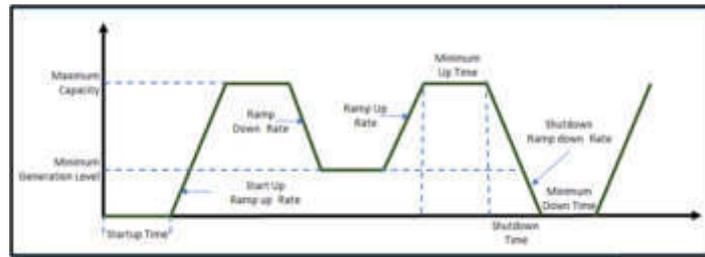


Fig. 3.1: System Ramping Capability (SRC)

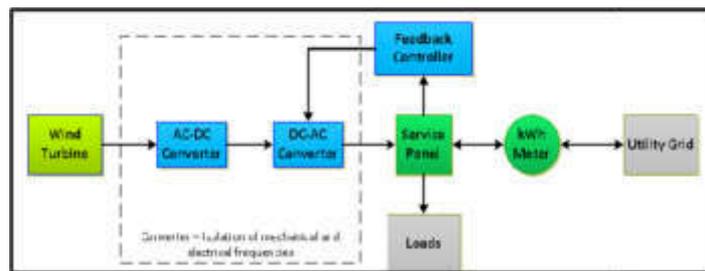


Fig. 3.2: Grid in energy management

3. To understand the order of frequency suitable for controlling the power system of wind farms
4. To evaluate the challenges of grid integration of wind power for wind farms

**3. Methodology.** The analysis of the grid integration comprises a vast course of practices with respect to the optimization of wind energy. The control of winds is an essential factor which required critical discussion [1]. The study can be obtained by secondary data analysis in terms of statistical and numerical evaluation. The analytical spectrum covers the maximum power point tracking, power quality issues and other such valid resources with respect to the issues related to the integrity of resources [2]. The use of qualitative data has been verified with valuable aspects providing an array of information related to the study.

**3.1. Role of grid electricity in managing both variability and uncertainty of loads.** The photovoltaic (PV) rays are the power system that serves well in managing both the variability and uncertainty of loads. Settings aside the practicality of the plan, the problem can be solved by reaching the global target by up to 20% significantly [3]. The electricity grid plays a crucial role in establishing the power station that results in the emission of power supplies.

*System Ramping Capability (SRC).* System ramping capability is the power that changes the direction of the wind by generating power. This speed is essential that can ramp up or ramp down based on the systematic supply functions [4]. Rating the ramp event is essential as it generates the maximum capacity of wind farms with a range smaller than 4h.

The above systematic diagram shows the quality of adaptability that provided a pathway towards a robust comprehensive declaration strategy [5]. The passage of electricity has been estimated to be productive in various ways.

*Purpose of the grid in energy management.* The link between the smart grid and energy management allows for monitoring the redundant factors with respect to preventing system disruptions. Electric grids in wind supply are the integrated solutions that are relevant to administration practices [7]. Smart grids use digital technologies, sensors and software applications which maintain the stability of the grid pertaining to the reliability of resources.

The figure 3.2 reflects the importance of the grid in energy management with a connection of the AC-DC

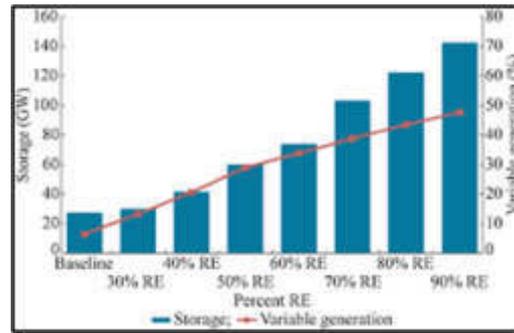


Fig. 3.3: Integrated variable renewable energy

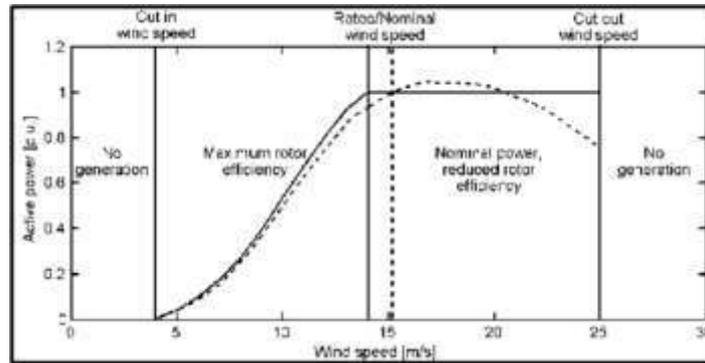


Fig. 3.4: Passive stall control

controller. The service practices are sufficient in the utility grid showing prominent outcomes.

*Integrated variable renewable energy.* Running the grid integration increases the generation and collection of loads that are metered to associate demand and supply services. An energy capacitor system has strong connectivity to stabilize the wind farms during a change in wind speed [8]. The voluntary use of resources has initiated the interconnections that seek to accommodate significant power system provisions.

The figure 3.3 highlights the systematic practices that are essential for fostering low emissions as well as the use of conventional energies. The percentage of storage increases with an increase in variable generations [9]. This has improved the quality of power stations by sharing the resources predominantly.

**3.2. Importance of generator modeling and wind turbine.** Wind energy plays a crucial role in establishing an environment-friendly low-carbon economy. Traditionally, DSC machines have synchronized the machines and induction tools that lead to the modeling of wind turbines [10].

*Passive stall control.* The utilization of wind energy is subjective to initializing passive stall control to pump water and to install various resources that create turbulence to the rotor blade that is designed to progress with a subtle process [11].

The figure 3.4 declares the passive role of practices that are relevant for medium and large-scale wind turbines, and inductors. These are needed to make permanent magnets that can consider various machines featuring guidelines for quicker reference [12].

*Rectifiers and inverters.* Rectifiers and inverters are the electrical mechanisms that aim in improving the electronic circuits for a better flow of current. They rectify the power supply and make the programming stable for wind farms [15].

The act of converting an alternative current into a direct current is relevant to the current flows that are

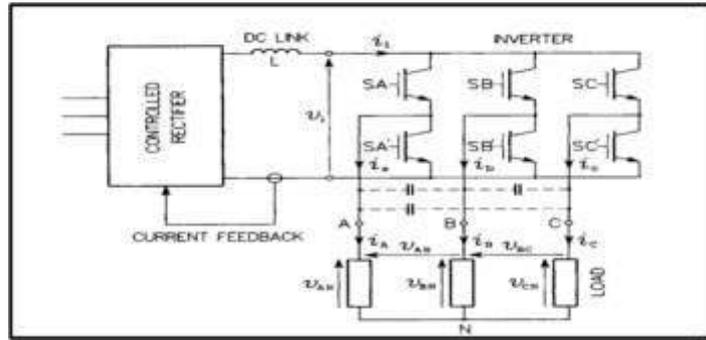


Fig. 3.5: Rectifiers and inverters

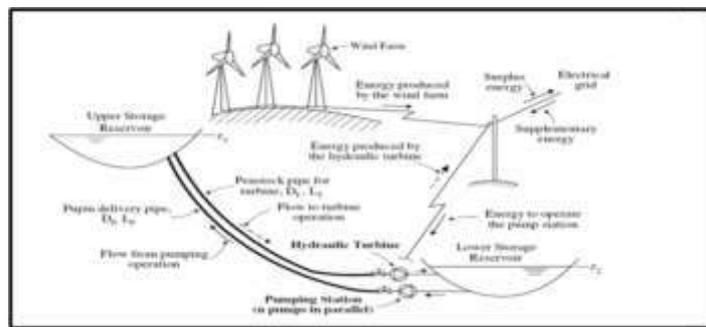


Fig. 3.6: Wind power integrated system

measured as rectification. Inverters increase the rate of sustainability by modifying the course of wind energy and with a better knowledge of controlled rectifiers.

**3.3. Order of frequency suitable for controlling the power system of wind farms.** The frequency level for controlling the power supply projects is managed by the transformer. When an alternative current is linked with alternating voltage, a magnetic field gets generated significantly [16].

*Wind power integrated system.* Grip integration is a collection of activities that are spilt into categories based on the activation of planning and execution activities. In addition to that, network development and system operations are relevant that help in increasing the power technology [17]. It helps in providing a significant change in the production system of power supply that allows proper optimization of wind energy in real-time analysis.

The figure 3.6 illustrates the terminologies based on energy production hydraulic turbines that are pumped up from the low storage reservoirs. With respect to the power delivery services, it is the generation of power supply that is objective in avoiding the unbalanced state of grid practices [18].

*Optimization of wind energy.* The optimization is obtained by associating a wind farm that helps in improving the feasibility of the wind energy system [19]. There are various ways odd improving wind turbine performance that can change the scenario towards sophistication.

1. Predicting the course of the wind
2. Establishing a constant speed transmitter
3. Variability in speed generator
4. Sophistication in loads programming and RPM

The systematic diagram elaborates on the features that increased the reliability of wind energy by withdrawing power from solar energy and converting it into power transmitters.

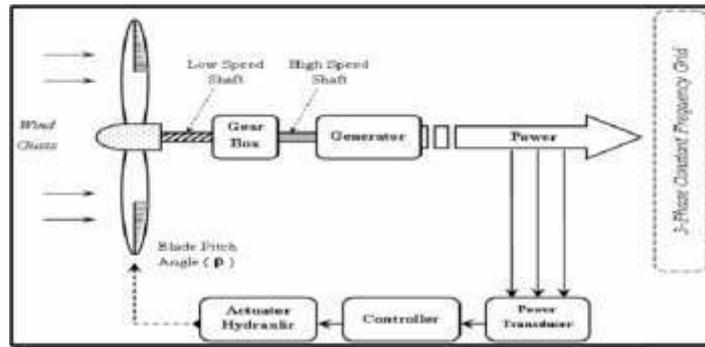


Fig. 3.7: Optimization of wind energy

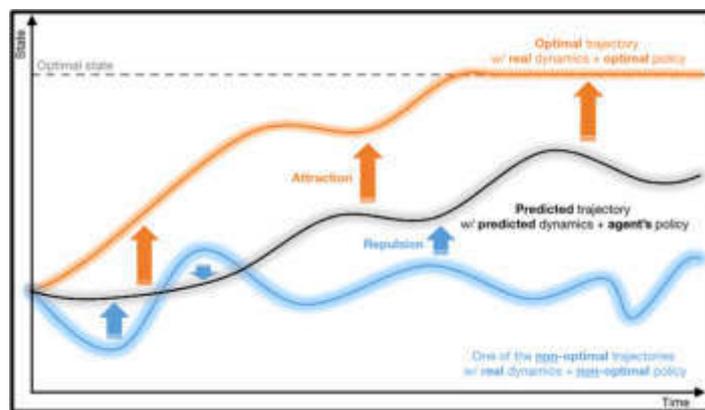


Fig. 3.8: A trajectory optimization model

*Trajectory optimization model.* The model is a power of sizing the trajectories for non-linear programming of wind energy which is a solution to all the negative consequences. Classically, the indirect methods have inculcated the optimization of partial and implicit differentiation in a significant manner [20].

The figure 3.8 highlights the process of attraction and repulsion that are interconnected in terms of non-optimal practices. Moreover, analyzing the wind scale and parameters of the trajectory optimization in terms of stabilizing the wind energy operations in a respective manner [21].

**4. Challenges of grid integration of wind power for wind farms.** One of the primary issues in grid integration is wind power is an intermittent nature of wind energy. The controlling elements are to be treated so as to generate turbine technology and protection issues [22]. Challenges address the instability and quality of energy integration that are outsourced as special elements.

*Output power prediction.* Although power prediction is an essential part of wind energy, it may sometimes give rise to various issues that include voltage fluctuations and the inability to meet reactants at the right time. This may lead to the deployment of resources which in turn is responsible for de-tracking reactive powers [23].

The high rate of fluctuation is clearly presented in the above figure 4.1 but fails to resonate with the variation in wind speed. This reduces the smoothness of variable speed up to 75% which in turn is responsible for the high consumption of power [24].

*Low voltage capability .* The ability of voltage protection reduces with excessive turbine programming resulting in disruptions in the power stations. The equivalence of the stimulation control is suppressed due to the post-fault segments which are currently irrelevant to develop non-superconducting issues [25]. The dispatchable issues are always high in the demand scale policies to a greater extent.

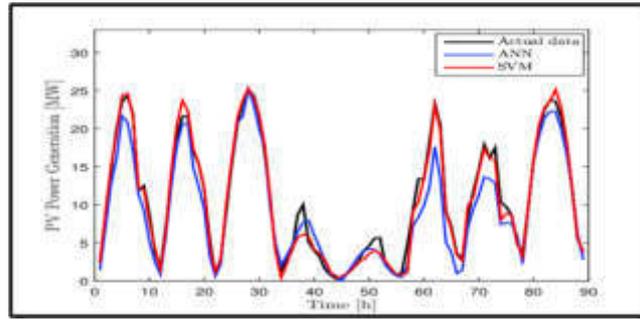


Fig. 4.1: Output power prediction

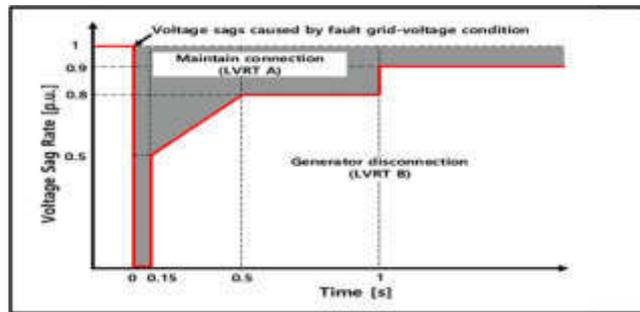


Fig. 4.2: Low voltage capability

The figure 4.2 illustrates the low voltage capability which is low key and is least maintained in the sage of voltage. It loses the capability to remain in service and to organize deep-volt segments [26].

*The issue in power quality.* The fault in the rising supply is detrimental in nature declining the efficiency level predominantly. The frequent processes may at times become unable in controlling the contingency events [27].

The figure 4.3 highlights that the rate of interruptions and imbalance in electrical flow is prevalent in nature. It can be proven that a deviation in frequency can cause harmful effects on the wind power supply in accordance with the harmonics of electricity.

**5. Results.** The above systematic diagram determines the factors that are respective to stabilizing power control. The ideas relevant to the practice of ideas are figured as original in terms of replenishing irrelevant prodigies.

*Power plant developers.* The range of power plant supplies induces the level of proximity in the field of electricity. Power plant developers have facilitated the course of building possibilities that are subjective to resource analysis [28].

*Information technology developers.* Developments are essential in terms of navigating the technical planings that are significant in working profitability.

*Service providers.* Service providers have a proper significance resolution of the fundamental approaches with the optimal supply level services. The range of possibilities rises with a rise in the association of significant applicants based on global programming of practices [29]. The verification is well established in accordance with the technical prospects based on projective modules to enhance the service provider practices. As a result of this, the relevant approaches have evoked the practice of pepper initiation in a significant manner.

*Variability and uncertainty of the power system.* When considering the power balance, the variability of the power system plays an important role that reducing the risk factors to a certain extent. Uncertainty exists in

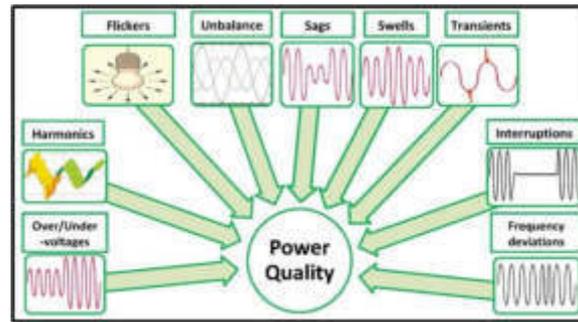


Fig. 4.3: Issue in power quality



Fig. 5.1: Factors responsible for conducting wind farms

measuring the forecast values that are respective of the multiple instances [6]. Variability is quantified by the distribution of power frequencies that are derived from sufficient data resources.

The structure shown in figure 5.2 proposes the separation of variability and uncertainty in the power system. The countermeasures address the efficiency of operations in applying the strategies that are determined to renewable energy resources significantly [6].

*Variable speed with partial scale frequency.* Variable speeds are an essential way of measuring the frequency of the scales that are used for converting kinetic energy with large power-capturing capabilities. Basically, partial scale frequency depends upon the level of electricity that are approachable in nature [13].

The systematic diagram shown in figure 5.3 focuses on the power rating scales that are defined as speed ranges. The smaller the frequency converter, the more conceptualized it is from an economic point of view [14].

**6. Conclusion.** In conclusion, the systematic examination presented in this study reveals a comprehensive understanding of the factors that play a pivotal role in stabilizing power control, particularly within the context of wind power stations. The novel insights into these factors are instrumental in optimizing the utilization of this global resource, ushering in real-time benefits and heightened electricity generation. The authenticity of these findings accentuates the significance of prioritizing effective power control, ultimately resulting in a

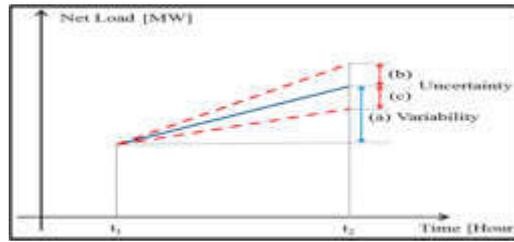


Fig. 5.2: Variability and uncertainty of power system

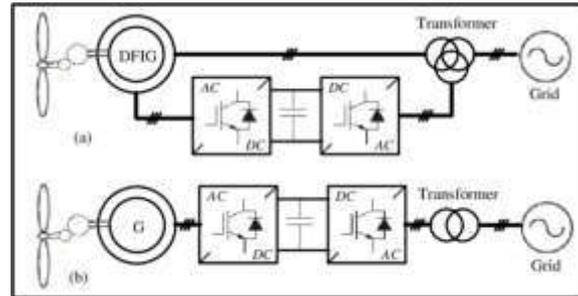


Fig. 5.3: Partial scale frequency level

heightened level of stability within grid-connected power supplies. The establishment of wind farms, as a direct outcome of this research, embodies a transformative step towards realizing enhanced reliability and efficiency in energy generation. This progression towards knowledge refinement underscores the importance of continually advancing the sophistication of practices in the field. By addressing the marginalization of resources and embracing innovative techniques, we can elevate the sustainability and dependability of wind power, further contributing to the broader goals of a sustainable energy future. The journey toward a greener and more resilient energy landscape is contingent upon the thoughtful application of such research-driven insights and the ongoing pursuit of excellence in power control within wind energy systems. The research may offer insights into power control and stability for specific wind energy setups, but generalizing these findings to all wind energy systems worldwide should be done with caution. Wind energy varies widely in terms of scale, technology, and environmental context, and the research may not account for all these nuances.

**Fundings.** This work is funded by the Science and Technology Project of SGCC (Research on Wind Generator Transient Voltage Optimal Control Technology in Weak Power Grid) (4000-202214067A-1-1-ZN).

#### REFERENCES

- [1] M. K. BAKSHIZADEH, S. GHOSH, Ł. KOCEWIAK, AND G. YANG, *Improved reduced-order model for pll instability investigations*, IEEE Access, (2023).
- [2] U. BURAGOHAIN AND N. SENROY, *Reduced order dfig models for pll-based grid synchronization stability assessment*, IEEE Transactions on Power Systems, (2022).
- [3] S. S. DAS AND J. KUMAR, *Socio-economic advancement and grid-frequency control of a wind farm-phs-solar connected ev hybrid electrical system considering imbalance prices*, Energy & Environment, (2023), p. 0958305X231192366.
- [4] W. DU, Y. WANG, H. WANG, B. REN, AND X. XIAO, *Small-disturbance stability limit of a grid-connected wind farm with pmsgs in the timescale of dc voltage dynamics*, IEEE Transactions on Power Systems, 36 (2020), pp. 2366–2379.
- [5] X. FENG, S. LIN, Y. LIANG, G. FAN, AND M. LIU, *Distributed risk-averse approximate dynamic programming algorithm for stochastic economic dispatch of power system with multiple offshore wind farms*, CSEE Journal of Power and Energy Systems, (2023).
- [6] C. GE, M. LIU, AND J. CHEN, *Modeling of direct-drive permanent magnet synchronous wind power generation system considering the power system analysis in multi-timescales*, Energies, 15 (2022), p. 7471.

- [7] M. HAJIAGHAPOUR-MOGHIMI, E. HAJIPOUR, K. A. HOSSEINI, M. TAVAKKOLI, S. FATTAHEIAN-DEHKORDI, M. VAKILIAN, AND M. LEHTONEN, *Hedging investments of grid-connected pv-bess in buildings using cryptocurrency mining: A case study in finland*, IEEE Access, (2023).
- [8] H. HUA, T. LIU, C. HE, L. NAN, H. ZENG, X. HU, AND B. CHE, *Day-ahead scheduling of power system with short-circuit current constraints considering transmission switching and wind generation*, IEEE Access, 9 (2021), pp. 110735–110745.
- [9] J. MA, Y. GU, Y. SHEN, Y. ZHOU, A. PHADKE, AND P. CHENG, *Stability analysis of power grid connected with direct-drive wind farm containing virtual inertia based on integrated dissipation energy model*, IEEE Transactions on Sustainable Energy, 12 (2021), pp. 2378–2392.
- [10] N. MOHAMMED AND M. CIOTARU, *Adaptive power control strategy for smart droop-based grid-connected inverters*, IEEE Transactions on Smart Grid, 13 (2022), pp. 2075–2085.
- [11] J. OUYANG, T. TANG, J. YAO, AND M. LI, *Active voltage control for dfig-based wind farm integrated power system by coordinating active and reactive powers under wind speed variations*, IEEE Transactions on Energy Conversion, 34 (2019), pp. 1504–1511.
- [12] M. S. PAULO, A. D. O. ALMEIDA, P. M. D. ALMEIDA, AND P. G. BARBOSA, *Control of an offshore wind farm considering grid-connected and stand-alone operation of a high-voltage direct current transmission system based on multilevel modular converters*, Energies, 16 (2023), p. 5891.
- [13] Z. PENG, Q. PENG, Y. ZHANG, H. HAN, Y. YIN, AND T. LIU, *Online inertia allocation for grid-connected renewable energy systems based on generic asf model under frequency nadir constraint*, IEEE Transactions on Power Systems, (2023).
- [14] A. PRAKASH, R. K. TIWARI, K. KUMAR, AND S. PARIDA, *Interacting multiple model strategy based adaptive wide-area damping controller design for wind farm embedded power system*, IEEE Transactions on Sustainable Energy, 14 (2022), pp. 962–973.
- [15] A. RAUF, M. KASSAS, AND M. KHALID, *Data-driven optimal battery storage sizing for grid-connected hybrid distributed generations considering solar and wind uncertainty*, Sustainability, 14 (2022), p. 11002.
- [16] B. SAHU AND B. P. PADHY, *Small signal stability analysis of dfig integrated power system considering pll dynamics under different grid strengths*, in 2023 IEEE PES Conference on Innovative Smart Grid Technologies-Middle East (ISGT Middle East), IEEE, 2023, pp. 1–5.
- [17] U. T. SALMAN, F. S. AL-ISMAIL, AND M. KHALID, *Optimal sizing of battery energy storage for grid-connected and isolated wind-penetrated microgrid*, IEEE Access, 8 (2020), pp. 91129–91138.
- [18] H. TAN, H. LI, R. YAO, Z. ZHOU, R. LIU, X. WANG, AND J. ZHENG, *Reactive-voltage coordinated control of offshore wind farm considering multiple optimization objectives*, International Journal of Electrical Power & Energy Systems, 136 (2022), p. 107602.
- [19] S. TAO, Q. XU, A. FELÍÓ, P. HOU, AND G. ZHENG, *Bi-hierarchy optimization of a wind farm considering environmental impact*, IEEE Transactions on Sustainable Energy, 11 (2020), pp. 2515–2524.
- [20] A. THAKALLAPPELLI, A. R. NAIR, B. D. BISWAS, AND S. KAMALASADAN, *Frequency regulation and control of grid-connected wind farms based on online reduced-order modeling and adaptive control*, IEEE Transactions on Industry Applications, 56 (2020), pp. 1980–1989.
- [21] M. N. UDDIN, N. REZAEI, AND M. S. ARIFIN, *Hybrid machine learning-based intelligent distance protection and control schemes with fault and zonal classification capabilities for grid-connected wind farms*, IEEE Transactions on Industry Applications, (2023).
- [22] X. WANG, J. ZHOU, B. QIN, AND L. GUO, *Coordinated power smoothing control strategy of multi-wind turbines and energy storage systems in wind farm based on madrl*, IEEE Transactions on Sustainable Energy, (2023).
- [23] T. WEN, Z. ZHANG, X. LIN, Z. LI, C. CHEN, AND Z. WANG, *Research on modeling and the operation strategy of a hydrogen-battery hybrid energy storage system for flexible wind farm grid-connection*, IEEE Access, 8 (2020), pp. 79347–79356.
- [24] Y.-K. WU, S.-M. CHANG, AND P. MANDAL, *Grid-connected wind power plants: A survey on the integration requirements in modern grid codes*, IEEE Transactions on Industry Applications, 55 (2019), pp. 5584–5593.
- [25] D. XIE, J. WANG, D. CHEN, C. DING, L. WANG, H. TIAN, AND S. LI, *A static equivalent method considering dfig grid-connected control strategy*, in International Conference on Energy and Environmental Science, Springer, 2023, pp. 231–244.
- [26] Y. YOO, S. SONG, J. SUH, J.-H. KIM, R.-J. PARK, AND S. JUNG, *Adaptive response method for communication failures of hierarchical reactive power control in wind farms*, IEEE Transactions on Sustainable Energy, 13 (2022), pp. 2343–2352.
- [27] Z. ZHANG AND X. ZHAO, *Coordinated power oscillation damping from a vsc-hvdc grid integrated with offshore wind farms: Using capacitors energy*, IEEE Transactions on Sustainable Energy, 14 (2022), pp. 751–762.
- [28] Z. ZHANG, M. ZHOU, Z. WU, S. LIU, Z. GUO, AND G. LI, *A frequency security constrained scheduling approach considering wind farm providing frequency support and reserve*, IEEE Transactions on Sustainable Energy, 13 (2022), pp. 1086–1100.
- [29] Y. ZHAO, K. XIE, C. SHAO, C. LIN, M. SHAHIDEPOUR, H.-M. TAI, B. HU, AND X. DU, *Integrated assessment of the reliability and frequency deviation risks in power systems considering the frequency regulation of dfig-based wind turbines*, IEEE Transactions on Sustainable Energy, (2023).

*Edited by:* Sathishkumar V E

*Special issue on:* Scalability and Sustainability in Distributed Sensor Networks

*Received:* Aug 25, 2023

*Accepted:* Oct 29, 2023



## RESEARCH ON SPEECH COMMUNICATION ENHANCEMENT OF ENGLISH WEB-BASED LEARNING PLATFORM BASED ON HUMAN-COMPUTER INTELLIGENT INTERACTION

YUFANG GU\*

**Abstract.** This study presents a novel web-based learning platform that leverages human-computer intelligent interaction to enhance English communication skills. The platform integrates cutting-edge technologies to create an immersive learning experience, combining natural language processing, speech recognition, and interactive exercises. Learners engage in real-time conversations with virtual tutors, receive personalized feedback, and access a vast repository of educational resources. The platform not only facilitates language acquisition but also encourages self-paced learning, making it a valuable tool for both educators and students. By harnessing the power of artificial intelligence, this web-based platform represents a significant advancement in the realm of English language education. To overcome these issues this paper proposed SVM with an improved satin Bower bird optimization algorithm (SVM-ISBBO). SVM-ISBBO uses fog computing services that minimize the latency and speeds up the process, effectively handling huge wearable devices. In this proposed work SVM-ISBBO monitors the students communication, vocal parameters, blood pressure, etc, and these values are obtained from wearable sensor devices and their notifications are sent back to teachers. Teachers diagnosed the student information and sent back the alert notifications to the students for taking proper medications. All this information is stored in fog-based cloud storage in a secure manner. The accuracy rate of KNN got 78.56%, NB got 81.74%, SVM got 85.15% and the proposed work of SVM-ISBBO got 92.34%.

**Key words:** English communication, fog computing, SVM, Satin Bower Bird Optimization, web communication.

**1. Introduction.** Nowadays, due to technological advancements in the communication field monitoring the health status of students and maintaining communication management. And the transmission of English communication information via a proper wireless communication network in a secure manner is a challenge. By implementing algorithms, handling unstructured data, and producing organized output in report format like electronic documents with reports in a secure[8]. Communication monitoring information about students is obtained from various wearable sensor devices on communication servers digitally. Due to the rapid growth of technology, various attackers can attack the transmission of communication information to clinical information[19]. For monitoring student information remotely based on IoT technology and collecting the smart health information of student's physical and mental strength, living environment, lifestyle, hereditary communication issues details, etc. This IoT-based communication information is transmitted to a cloud storage computing service through the wireless network. from the cloud storage, teachers can monitor and diagnose students' health conditions and send notification messages like early prevention of disease and its corresponding treatment [14].

The communication cyber-physical system uses smart communication sensor devices, a wireless communication network, and managing the sensor signals generated from communication sensor devices. To implement the MCPS, various machine learning and deep learning algorithm are applied in the monitoring of communication English communication information[17]. Cyber-physical system detects attacks in the storage of communication information in cloud computing. The main objective of the CPS environment is providing the secure using an artificial intelligence-based environment [23].

L. Zheng et.al[24] presented a system for producing an alert signal at the early stage warning for high-risk students using deep learning algorithms and maintaining electronic health records. Shuwandy et al. [15] described that CPS embedded the sensor devices with IoT technology and stored the communication information in the cloud storage platform. Ahmad Ali AlZubi et.al [4] proposed that Cyber-attack detection in communication uses cyber-physical systems and machine learning techniques. Wazid et al[20] propose AI-based technology

---

\*School of Foreign Languages Hubei Engineering University Xiaogan 432000 Hubei School of Graduate Studies Baliuag University Baliuag 3006, Bulacan, Philippines ([guyufang@hbeu.edu.cn](mailto:guyufang@hbeu.edu.cn))

for the prediction of a cybercrime attack in communication information and it uses an edge-based IoT environment. In view of existing research work, the issues are inefficient in handling huge wearable sensor devices, and lack of detection of the attack in the storing of English communication information. To overcome these issues this paper proposed SVM with an improved satin Bower bird optimization algorithm (SVM-ISBBO). SVM-ISBBO uses fog computing services that minimize the latency and speeds up the process, effectively handling huge wearable devices. The main contribution of this work is:

1. Monitoring student communication level these values are obtained from wearable sensor devices and stored in a fog node-based cloud storage platform.
2. For monitoring the abnormal condition and providing secure storage of student health information in a fog node-based cloud storage platform by implementing SVM with the improved satin bowerbird optimization algorithm.

The paper has been organized as follows: Section 2 describes the review of the literature, Section 3 methodology of monitoring and secure storage of communication cyber-physical system-based student English communication information using SVM-ISBBO, Section 4 discusses the experimented results and Section 5 concludes the paper with future directions.

**2. Review of Literature.** In the IoT-based technology of cloud computing embedded computation resources, heterogeneous data structure, storage capacities, and high-speed processing have been included in the development of smart communication systems. This fog computing links the smart sensor devices with the cloud storage platform. It provides a quick response time, but the issue with cloud computing is the negative impact of handling real-time reactions[2]. Fog computing is used in the English communication system which includes three levels of processing namely data is collected from edge devices of sensors, multiple devices are interlinked with one another, and processing of collected sensor data takes less than a second with its decision-making process[16].

The smart English communication system has become popular in the earlier diagnosis of a disease which provides proper treatment to the student. At the same time, a smart communication system consists of various sensors with MRI, PET, CT, etc. In order to provide the best quality of treatment to the student as well as to control the spread of the virus smart communication system is required. This smart communication system is monitored by a web based monitoring system[3]. A. A. Mutlag et.al[12] described the concept of enabling the technologies for fog computing in communication IoT systems. Kontakt LS, et al [10] presented an accurate prediction of the speech issues of the student by applying the machine learning technique and securely storing health information in a cloud storage platform.Li , et al[11] described the handling of communication information from various communication sensor devices and analysis stored in cloud storage platforms as well as handling big data analytics.Wu et al [21] presented that cyber-physical systems are based on preserving and securely storing English communication information and also detecting malicious attacks in the CPS. Thamilarasu et al [18] proposed that detection of attack in the English communication information is stored in cloud computing and also the detection of attacks in IoMT.

**3. Methodology.** An effective communication cyber-physical system for monitoring and securely storing communication information using fog computing service based on SVM with improved satin bower bird optimization algorithm (SVM-ISBBO). This communication cyber-physical system integrates the various communication sensor devices and obtained the communication-based sensor information and stores it in a fog node-based cloud storage system in a secure manner. The framework of SVM-ISPBO is given in Figure 3.1.

Figure 3.1 contains three layers data collection layer, the communication data stored in fog node-based cloud storage layer, and the communication data information processing layer.

**3.1. Data Collection Layer.** From the student, communication sensor signal information is collected and stored in a fog node-based cloud storage system. This collected communication information is diagnosed by the teachers and updates the current health status of students via a mobile app. The wearable sensor devices which are used in the monitoring of the health status of the students are temperature sensors, blood pressure sensors, heartbeat sensors, ECG sensors, EEG sensors etc. This sensor information is collected via wireless communication. These sensor devices are embedded with ESP8266 Node MCU microcontroller[6]. It is high processing speed, low cost, and provides accurate results.

Table 2.1: Survey on the smart communication system in CPS

<i>Author Name.</i>	<i>Sensors used.</i>	<i>Description.</i>
Ahmed Elhadad et.al (2022) [1]	temperature, ECG, and blood pressure sensors	Communication Monitoring System for Managing the Real-Time Notification in a fog computing system.
S. Senganet.al (2022) [13]	temperature sensor, Heartbeat sensor	Smart communication communication system for providing security devices on MIoT using Raspberry Pi
I. Ahmed et.al (2021) [1]	temperature sensor, Heartbeat sensor	A deep-learning-based smart communication system for student discomfort detection at the edge of the Internet of Things
Al-Sheikh et.al (2020) [2]	Heart rate, ECG, and body temperature sensor	Arduino, NodeMCU is used for transferring information via Wi-Fi.
Elango et.al (2020) [7]	Heartbeat sensor, body temperature sensor,	Storing health information in NodeMCU edge devices and implementing the communication protocols of Wi-Fi/ HTTP, and MQTT.
Islam et al. (2020) [9]	Heartbeat sensor, body temperature sensor,	Storing health information in NodeMCU edge devices and implementing the communication protocols of Wi-Fi/ HTTP, and MQTT.

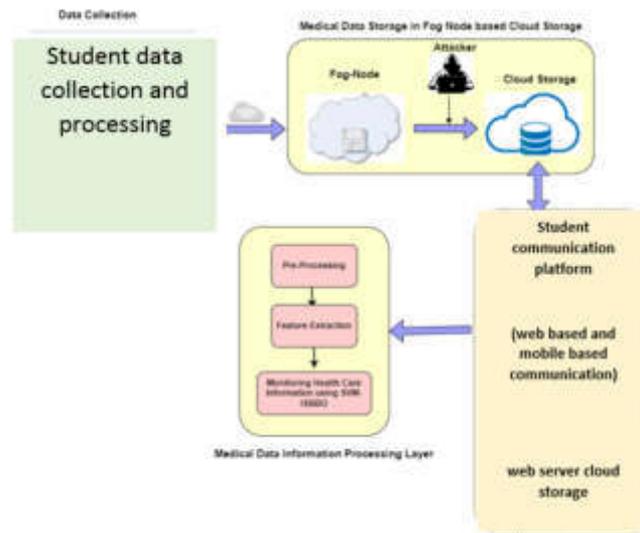


Fig. 3.1: Framework of SVM-ISPBO

**3.2. Communication Data Storage in Fog-based Cloud Storage.** The fog node lies between communication sensor devices and cloud storage computing. Sensor communication data were obtained from various wearable sensor devices. It minimizes latency, provides security, enhances the reliability of storing information, and has high storage capacity. Communication data were collected from all sensor devices and it has been transmitted to fog-node-based cloud storage. These communication data are transmitted for a pre-processing stage of the communication data information processing layer. In the monitoring of communication information, fog computing facilitates the processing of communication data without reducing its dimensionality of data and also prevents congestion of the network. Communication data stored in the cloud storage are transmitted to a communication server, the teachers monitor the health information and send the health status to the student via mobile application.

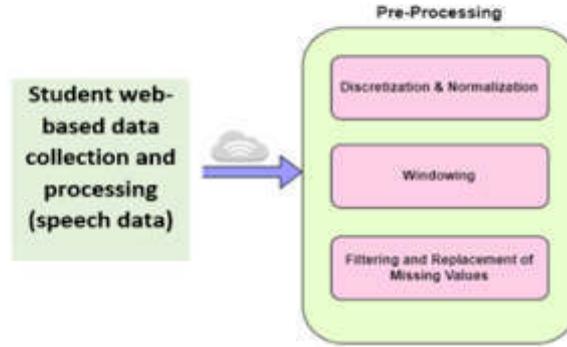


Fig. 3.2: Pre-Processing

**3.3. Communication Data Information Processing Layer.** The communication data information processing layer contains three modules namely pre-processing, feature extraction, and monitoring English communication information using SVM-ISPBO.

**3.3.1. Pre-Processing.** The communication sensor information is obtained from various sensor devices and it is in various formats like images, time series, numerical values, etc. And also, it may contain random samples as well as continuous sample structures. Therefore, pre-processing steps are needed for handling various formats of collected data. The steps involved in the pre-processing are given in Figure 3.2.

*Discretization & Normalization.* For the discretization of collected communication sensor data, the aspects of frequency and time are essential one. For that this paper uses standardized collected sensor values in the range between 0 and 1 is termed normalization. In order to handle the complex variations of sensor data and for the prediction of diseases based on the regression analysis and it can be evaluated as:

$$Q = \beta_0 + \beta_i M + \epsilon_i \text{ for } i = 1, 2, \dots, m \quad (3.1)$$

The sample input sensor communication data along with its variance and error value is termed by  $\epsilon_i$ . The least square values are denoted by  $\beta_0$  and  $\beta_i$ . The average values are evaluated from the sample input communication sensor data, by using standard deviation which is defined as:

$$\mu = \frac{\sum_{i=1}^n M_i}{\text{Frequ}} \quad (3.2)$$

Here, *Frequ* represents the frequency of data and  $M_i$  is the sample communication input sensor data. Then the normalization is denoted by:

$$n_q = \frac{\epsilon_i^*}{\sigma_i} \quad (3.3)$$

$$n_q = \frac{M_i - \mu_i}{\sigma_i} \quad (3.4)$$

Here  $\epsilon_i^*$  is residual value and  $\sigma_i$  is variance.

*Windowing.* The basic concept of windowing is splitting of communication sensor signal values into small segments in the aspect of time domain. Here the communication input sequence of sensor signals is  $sen_1, sen_2, \dots, sen_n$  and it is separated into windows with equal number of sensor activities  $acw_1, acw_2, \dots, acw_n$ . Here  $acw_1$  is defined as window by signifying  $[sen_i - \Delta sen, sen_i]$ . The length of the window is varied from one window to another. Based on the time domain two sensor signal activities are fall into the same window.

**3.3.2. Feature Extraction.** In order to detect the attack in the storing of communication information feature extraction is an essential one. It extracts the significant features or relevant features from the dataset. In this work of SVM-ISPBO, which extracts the statistical features of mean, entropy, variance, and standard deviation.

*Mean.* It is the ratio of sum of communication data to the total number of communication data in the database. It is represented as:

$$mean = \frac{1}{y} \sum_{i=1}^y Q_i \quad (3.5)$$

Here  $y$  represents the total instances in the dataset and  $Q_i$  is the  $i^{th}$  data.

*Entropy.* Analysing the probability of arranging communication sensor data with  $n$  ways and it is defined as:

$$entropy = - \sum_{x=1}^Q p_x \log_2 p_x \quad (3.6)$$

Here,  $p_x$  is the probability of data and  $x$  is the possible sample data values.

*Variance.* It is the mean squared difference between every communication sensor data in the dataset and its mean value, which is represented by:

$$variance = \frac{1}{y-1} \sum_{i=1}^y (Q_i - mean)^2 \quad (3.7)$$

*Standard Deviation.* It is defined as root of variance, which is represented as follows:

$$\sqrt{\frac{1}{y-1} \sum_{i=1}^y (Q_i - mean)^2} \quad (3.8)$$

### 3.3.3. Monitoring MCPS based Communication information using SVM-ISBBO (Proposed).

In the communication cyber physical system-based model monitoring the student information and storing the information securely in the fog node-based cloud storage system uses support vector machine with optimized algorithm of improved satin bower bird optimization algorithm (SVM-ISBBO). In this communication cyber physical system are integrates the various communication sensor devices and obtained the communication-based sensor information and stored in fog node-based cloud storage system in a secure manner.

**3.3.4. Support Vector Machine in MCPS web communication monitoring.** In the smart English communication cyber physical system, communication sensor signals are obtained from various communication sensor devices. After applying the pre-processing stage, extract the relevant features of the communication sensor information. SVM algorithm is used for analysis of collected communication sensor signals of the student. By using MCPS, SVM algorithm classifies the communication sensor information based on its range of value. MCPS monitored the communication sensor signal information and analysis that critical situation of student, immediately itsends the alert message to student's physician and caretaker.

The mathematical formulation of applying SVM is to find the optimal hyperplane which separates the features of the labelled data. Let  $\mathcal{H}$  is the Hilbert space with inner product of  $\langle \cdot, \cdot \rangle$  and its induced norm  $\| \cdot \|$ . let us consider  $\mathbb{R}^l$  be the m-dimensional Euclidean distance space and it is defined as  $\emptyset : \mathbb{R}^l \rightarrow \mathcal{H}$  is a mapping function. For a training communication data set,

$$MTD_l = \{(p_i, q_i) | p_i \in \{-1, 1\}\}_{i=1}^l \quad (3.9)$$

Here  $q_i$  is the label of  $p_i$  and its marginal function  $maf_l$  is defined as follows:

$$maf_l(p) = \langle \omega, \emptyset(p) \rangle + \theta, p \in \mathbb{R}^l \quad (3.10)$$

Here  $\theta$  is the bias term, to yield optimal hyperplane  $\omega$  and  $\theta$  are used in training data set. To get the optimal hyperplane, objective function is required,

$$\begin{aligned} \text{Minimize: } t(\omega, \theta) &= \frac{1}{2} \|\omega\|^2 + Cp \sum_{i=1}^k \xi_i \\ \text{Subject to constraints: } q_i(\omega, \theta(a_i) + \theta) &\geq 1 - \xi \end{aligned}$$

$$\xi \geq 0, \quad i = 1, 2, \dots, k$$

Here  $Cp$  is the regularization parameter and by using Lagrange multiplier method Equation 3.3.4 can be rewritten as follows to get optimal vector  $\omega$ .

$$\omega = \sum_{i=1}^l \alpha_i q_i \phi(p_i)$$

Subject to constraints :  $\sum_{i=1}^l \alpha_i q_i = 0$

Here  $\alpha_i$  is  $i = 1, 2, \dots, l$ . Substitute Equation 3.3.4 in Equation 3.9,

$$\begin{aligned} maf_l(p_0) &= \langle \omega, \phi(p_0) \rangle + \theta \\ &= \left\langle \sum_{i=1}^l \alpha_i q_i \phi(p_i), \phi(p_0) \right\rangle + \theta \\ &= \sum_{i=1}^l \alpha_i q_i \langle \phi(p_i), \phi(p_0) \rangle + \theta \end{aligned} \quad (3.11)$$

Let  $L : \mathbb{R}^l \times \mathbb{R}^l \rightarrow \mathbb{R}$  be a kernel function and it can be expressed as follows:

$$L(p_i, p_j) = \langle \phi(p_i), \phi(p_j) \rangle \quad (3.12)$$

Then based on Equation 3.11

$$mtd_k(p_0) = \sum_{i=1}^l \alpha_i q_i L(p_i, p_0) + \theta \quad (3.13)$$

Now SVM can be represented as follows:

$$sv_l(p_0) = \text{sgn}(maf_l(p_0)) = \begin{cases} 1 & \text{if } maf_l(p_0) \geq 0 \\ -1 & \text{if } maf_l(p_0) < 0 \end{cases} \quad (3.14)$$

Here the support vectors are trained with training data set of non-zero value of  $\alpha$ . The kernel function  $L$  and coefficient  $C$ , and it is used to classify the data. This SVM technique monitors the abnormal status of student health status and securely storing of communication information in fog node-based cloud storage platform. To improvising and more accurate monitoring of abnormal status of student health status based on the wearable sensor information optimization algorithm is required. Therefore, this work implements the improved satin bowerbird optimization algorithm.

**3.3.5. Improved Satin Bower Bird Optimization Algorithm.** Satin Bowerbird Optimizer (SBBO) is an intelligent optimization algorithm, determines that in the wild the adult male of Satin Bowerbird which simulates the breeding behavior. It has a strong power of survival and skill in reproduction. By constructing courtship cabin the matured Satin Bowerbird wins the female by attracting it by holding aluminous object in its beak, continuous loud singing which improves the probability of successful courtship. The male Satin Bowerbird constantly resists the challenges of its competitors and protect its nest from damage based on its survival rules. The steps involved in the SBO for the detection of attack in MCPS model and monitors the student information in an effectively.

*Step 1.* For initialize the population of Satin Bowerbirds randomly, this paper uses Logistic chaotic map for initialize the population of Satin Bowerbirds which improves the diversity of initial population, optimized accuracy and speed in convergence. Therefore, it is called as improved Satin Bowerbird optimization algorithm (ISBBO). This Logistic chaotic map is defined as:

$$P_{i+1} = \mu P_i * (1 - P_i) \tag{3.15}$$

The value of  $\mu$  as a control parameter which ranges from 0 to 5. When the value of  $\mu$  is larger as 5, then its initialization of population will get enhanced. In the solution space various courtship cabin is generated for various satin bowerbird  $MB$ .

*Step 2.* Evaluate the fitness function of each individual satin bowerbird and compute the ratio of fitness function to overall fitness function value based on the probability of individual selection of satin bowerbird. This probability of selection of courtship cabin is computed by:

$$Proba_j = \frac{fit_j}{\sum_{m=1}^{MB} fit_m} \tag{3.16}$$

$$fitn_j = \begin{cases} \frac{1}{1+f(y_j)} , & f(y_j) \geq 0 \\ 1 + \lceil f(y_j) \rceil , & f(y_j) < 0 \end{cases} \tag{3.17}$$

The fitness value of  $fitn_j$  represents that  $j^{th}$  courtship cabin and  $f(y_j)$  is the objective function of  $j^{th}$  courtship cabin

*Step 3.* Update the position of male Satin Bowerbird optimization algorithm in the last iteration by using:

$$y_{jk}^{iter+1} = y_{jk}^{iter} + \lambda_k \left( \left( \frac{y_{lk} + y_{elite,k}}{2} \right) - y_{jk}^k \right) \tag{3.18}$$

Here  $y_{jk}^{iter}$  is the k-dimensional component of  $j^{th}$  individual male satin bowerbird in the  $iter^{th}$  iteration.  $y_{lk}$  is the k-dimensional component of present global optimal position of population. The step factor of  $\lambda_k$  is evaluated by:

$$\lambda_k = \frac{\alpha}{1 + a_j} \tag{3.19}$$

Here  $\alpha$  is the maximum step size and  $a_j$  is the probability of choosing the target courtship cabin with value [0,1]. From the Equation 3.19 clearly shows that for the selection of target location with respect to greater probability along with its smaller step size. If the probability of selection of target location is 0 means, then its step size is largest and it is represented as  $\alpha$ . Similarly, If the probability of selection of target location is 1 means, then its step size is largest and it is represented as  $\alpha/2$ .

*Step 4.* To prevent local optimization, strong male satin bowerbird often steals from other male satin bowerbird courtship cabins and destroys the cabins. At the end of each iteration of ISBBO algorithm, randomly improve the mutation process. At this stage,  $y_{jk}$  follows the normal distribution and it is represented as:

$$y_{jk}^{iter+1} \sim M(y_{jk}^{iter}, \sigma^2) \tag{3.20}$$

$$M(y_{jk}^{iter}, \sigma^2) = y_{jk}^{iter} + (\sigma * M(0, 1)) \tag{3.21}$$

The evaluation of standard deviation is:

$$\sigma = z * (vari_{max} - vari_{min}) \tag{3.22}$$

Here, scaling factor is denoted by  $z$  and  $vari_{max}$ ,  $vari_{min}$  are upper limit and lower limit of variable  $y_i$ .

Table 4.1: Description of features in the Cleveland dataset

Features	Description
p.Age	Age of student in years
p.Sex	Female or Male
p.Chole	Measure of serum cholesterol
p.Chp	Types of chest pain
T_restbps	Resting blood pressure
F_bs	fasting blood sugar
Rest_ecg	Resting measure of electrocardiographic
Thalach	maximum heart rate
Exang	Exercise caused angina
Oldpeak_ST	ST depression caused by exerciserelative to rest
Slope	Slope of the peak exercise ST segment
Ca	Number of major vessels of coloured by Flourosopy
Thal	Type of Defect

*Step 5.* Based on initial population and population generated from mutation process in each iteration new population is formed. The fitness value function is generated for all individual male satin bower bird with its combination of population. This generated population is arranged in ascending order. In the analysis of individual male satin bowerbird has minimum fitness function value are removed. It retained only the largest fitness function value of male satin bowerbird. In this stage the optimal solution is computed and repeat the process until it reaches the maximum iteration. This SVM-ISBBO algorithm improvising and produce more accurate monitoring of abnormal status of student health status based on the wearable sensor information and send alert message to the student through physician in proper time. The student information also securely stored in the fog node based cloud storage platform.

**4. Result & Discussion.** The proposed work of SVM-ISBBO is used for monitoring the health information and securely storing the communication information. This proposed work SVM-ISBBO is evaluated the performance of analysis in the aspect of detection of accuracy, prediction of attack, ratio of delay, cost of communication, sensitivity, specificity and F1-Score.

**4.1. Data Set Description.** The dataset used in an effective communication cyber physical system for monitoring and securely storing communication information using fog computing service is Cleveland dataset from the UCI repository. The attributes are collected from various communication sensor devices like temperature sensor, ECG, pulse oximeter, temperature, and blood pressure sensors. The data were collected and saved in the fog node-based cloud storage platform. It consists of 5060 records, 13 features. Table 4.1 is the description of Cleveland dataset. This dataset contains heart disease related information.

This proposed work is compared with existing algorithms of KNN [22], NB, SVM[5].

*True Positive Rate (TPR).* The attack in the MCPS is exactly classified as an attack is known as sensitivity.

$$TPR = \frac{TP}{TP+FN} \quad (4.1)$$

*False Positive Rate (FPR) or Miss Rate.* It is also called as false alarm. This identified the normal data is considered as an incorrectly identified as an attack.

$$FPR = \frac{FP}{FP+TN} \quad (4.2)$$

*False Negative Rate (FNR) or Fall Out.* It is also called as miss rate. Incorrectly identified the attack data. that is, true positive will be missed.

$$FNR = \frac{FN}{FN+TP} \quad (4.3)$$

Table 4.2: performance metric measures of SVM-ISBBO

Algorithm	TPR	TNR	Miss Rate	Fall Out	MCC
KNN	0.793	0.764	0.078	0.061	0.25
NB	0.789	0.752	0.082	0.064	0.36
SVM	0.863	0.834	0.087	0.59	0.21
SVM-IBBO(Proposed)	0.931	0.911	0.059	0.051	0.18

*True Negative Rate (TNR).* The normal data is correctly predicted as normal and it is called as specificity.

$$TNR = \frac{TN}{TN+FP} \tag{4.4}$$

*Mathews Correlation Coefficient (MCC).* It is a corelation between predicted output with real data.

$$MCC = \frac{(TP.TN) - (FP.FN)}{\sqrt{(TP+FP).(TP+FN).(TN+FP).(TN+FN)}} \tag{4.5}$$

*MSE.* The mean squared error (MSE) calculate the average of the squares of the differences between the predicted values and actual values.

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_{pi} - y_{ai})^2 \tag{4.6}$$

*RMSE.* It is similar to MSE but for compute this by root of MSE.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_{pi} - y_{ai})^2} \tag{4.7}$$

From the table 4.2, the proposed SVM-IBBO in the monitoring of health information using MCPSis given. The proposed work of SVM-IBBO got 0.931 in TPR, 0.911 in TNR, 0.059 for miss rate , 0.051 in fall out and MCC 0.18.

*Percentage Rate of Delivery Communication Sensor Data Packet.* The percentage of delivery of communication sensor information is determined by number of packets of communication information delivered for a particular time in an effective way. The percentage between the sent packet of communication sensor data and obtained packet of communication sensor data is evaluated as delivery percentage rate of communication sensor data packets. This can be shown in Figure 4.1.

From the Figure 4.1 it seems that the number of sensor devices increases and transmitting rate of communication sensor values are decreased in the proposed work of SVM-IBBO. Therefore, it produces high delivery percentage rate of communication sensor data packets.

*Response Rate.* When transmission of communication sensor information from various sensor devices to fog node-based cloud storage platform. Response rate is rate of interaction of sensor devices for obtained the sensor information and stored it in fog node. This period of response in the network is given in Figure 4.2.

From the Figure 4.2 seems that the proposed work SVM-IBBO gives high response rate of transmission of communication sensor information. Figure 4.3 shows that error rate in the monitoring of English communication sensor information.

In the analysis of monitoring of health-based sensor information our proposed work SVM-ISBBO requires minimum error rate. Figure 4.4 shows that accuracy rate of proposed work.

In the analysis of Figure 4.4, shows that accuracy rate of KNN got 78.56% , NB got 81.74%, SVM got 85.15% and proposed work of SVM-ISBBO got 92.34%.

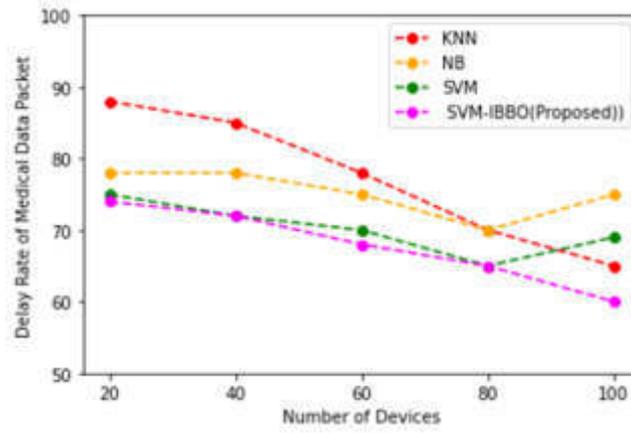


Fig. 4.1: Percentage Rate of Delivery Communication Sensor Data Packet

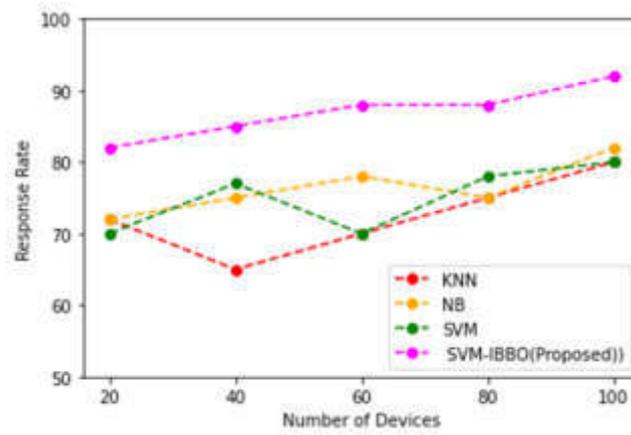


Fig. 4.2: Response Rate

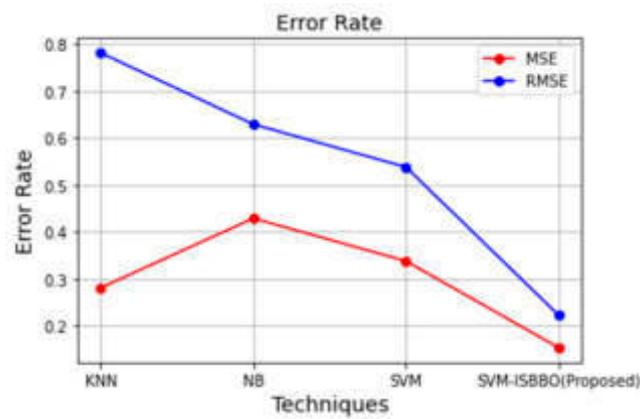


Fig. 4.3: Error Rate

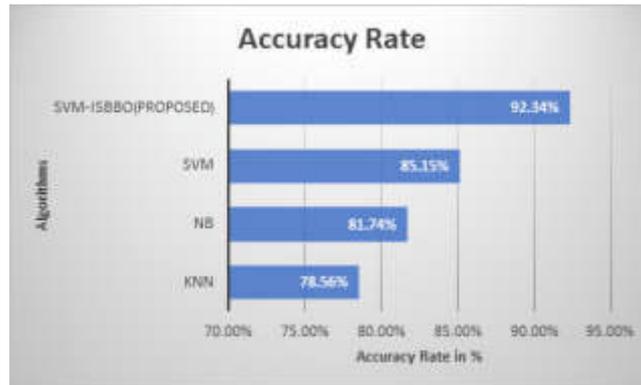


Fig. 4.4: Accuracy Rate

**5. Conclusion.** An effective communication cyber physical system for monitoring and securely storing communication information using fog computing service using SVM-ISBBO was implemented. Communication sensor information is obtained from various wearable sensor devices of student's body and storing this information securely in fog node computing until it safely transmitted to communication server for permit the teachers' heart rate, temperature, blood pressure etc and these values are obtained from wearable sensor devices and its notifications are sent back to teachers. Teachers' diagnosis the student information and sent back the alert notifications to the students for taking proper medications. The accuracy rate of KNN got 78.56% , NB got 81.74%, SVM got 85.15% and proposed work of SVM-ISBBO got 92.34%. This proposed work produces high performance in various aspects of accuracy, TPR, TNR, Miss rate, fall out and monitor the abnormal communication sensor information in an efficient and accurately. In future, this work may be extended up to for providing more secure in the fog node by applying the encryption of communication sensor data.

## REFERENCES

- [1] I. AHMED, G. JEON, AND F. PICCIALLI, *A deep-learning-based smart healthcare system for patient's discomfort detection at the edge of internet of things*, IEEE Internet of Things Journal, 8 (2021), pp. 10318–10326.
- [2] M. A. AL-SHEIKH AND I. A. AMEEN, *Design of mobile healthcare monitoring system using iot technology and cloud computing*, in IOP conference series: materials science and engineering, vol. 881, IOP Publishing, 2020, p. 012113.
- [3] A. ALDAHRI, B. ALRASHED, AND W. HUSSAIN, *Trends in using iot with machine learning in health prediction system*, Forecasting, 3 (2021), pp. 181–206.
- [4] A. A. ALZUBI, M. AL-MAITAH, AND A. ALARIFI, *Cyber-attack detection in healthcare using cyber-physical system and machine learning techniques*, Soft Computing, 25 (2021), pp. 12319–12332.
- [5] N. M. J. AUGUSTINE AND S. R. N. SAMY, *Smart healthcare monitoring system using support vector machine*, Australian Journal of Science and Technology, 2 (2018), pp. 1–8.
- [6] N. DOCUMENTATION, *Nodemcu documentation*, URL: [nodemcu.readthedocs.io/en/master/](http://nodemcu.readthedocs.io/en/master/), (2019).
- [7] K. ELANGO AND K. MUNIANDI, *A low-cost wearable remote healthcare monitoring system*, Role of Edge Analytics in Sustainable Smart City Development: Challenges and Solutions, (2020), pp. 219–242.
- [8] T. M. GHAZAL, *Internet of things with artificial intelligence for health care security*, Arabian Journal for Science and Engineering, (2021).
- [9] M. M. ISLAM, A. RAHAMAN, AND M. R. ISLAM, *Development of smart healthcare monitoring system in iot environment*, SN computer science, 1 (2020), pp. 1–11.
- [10] L. S. KONDAKA, M. THENMOZHI, K. VIJAYAKUMAR, AND R. KOHLI, *An intensive healthcare monitoring paradigm by using iot based machine learning strategies*, Multimedia Tools and Applications, 81 (2022), pp. 36891–36905.
- [11] W. LI, Y. CHAI, F. KHAN, S. R. U. JAN, S. VERMA, V. G. MENON, F. KAVITA, AND X. LI, *A comprehensive survey on machine learning-based big data analytics for iot-enabled smart healthcare system*, Mobile networks and applications, 26 (2021), pp. 234–252.
- [12] A. A. MUTLAG, M. K. ABD GHANI, N. A. ARUNKUMAR, M. A. MOHAMMED, AND O. MOHD, *Enabling technologies for fog computing in healthcare iot systems*, Future generation computer systems, 90 (2019), pp. 62–78.
- [13] S. SENGAN, O. I. KHALAF, S. PRIYADARSINI, D. K. SHARMA, K. AMARENDRA, AND A. A. HAMAD, *Smart healthcare security*

- device on medical iot using raspberry pi*, International Journal of Reliable and Quality E-Healthcare (IJRQEH), 11 (2022), pp. 1–11.
- [14] S. SENGAN, O. I. KHALAF, G. R. K. RAO, D. K. SHARMA, K. AMARENDRA, AND A. A. HAMAD, *Security-aware routing on wireless communication for e-health records monitoring using machine learning*, International Journal of Reliable and Quality E-Healthcare (IJRQEH), 11 (2022), pp. 1–10.
- [15] M. L. SHUWANDY, B. ZAIDAN, A. ZAIDAN, A. S. ALBAHRI, A. H. ALAMOUDI, O. S. ALBAHRI, AND M. ALAZAB, *mhealth authentication approach based 3d touchscreen and microphone sensors for real-time remote healthcare monitoring system: comprehensive review, open issues and methodological aspects*, Computer Science Review, 38 (2020), p. 100300.
- [16] A. I. TALOBA, R. ALANAZI, O. R. SHAHIN, A. ELHADAD, A. ABOZEID, A. EL-AZIZ, M. RASHA, ET AL., *Machine algorithm for heartbeat monitoring and arrhythmia detection based on ecg systems*, Computational Intelligence and Neuroscience, 2021 (2021).
- [17] S. TEPJIT, I. HORVÁTH, AND Z. RUSÁK, *The state of framework development for implementing reasoning mechanisms in smart cyber-physical systems: A literature review*, Journal of computational design and engineering, 6 (2019), pp. 527–541.
- [18] G. THAMILARASU, A. ODESILE, AND A. HOANG, *An intrusion detection system for internet of medical things*, IEEE Access, 8 (2020), pp. 181560–181576.
- [19] A. TIWARI, V. DHIMAN, M. A. IESA, H. ALSARHAN, A. MEHBODNIYA, M. SHABAZ, ET AL., *Patient behavioral analysis with smart healthcare and iot*, Behavioural Neurology, 2021 (2021).
- [20] M. WAZID, P. RESHMA DSOUZA, A. K. DAS, V. BHAT K, N. KUMAR, AND J. J. RODRIGUES, *Rad-ei: A routing attack detection scheme for edge-based internet of things environment*, International Journal of Communication Systems, 32 (2019), p. e4024.
- [21] D. WU, H. ZHU, Y. ZHU, V. CHANG, C. HE, C.-H. HSU, H. WANG, S. FENG, L. TIAN, AND Z. HUANG, *Anomaly detection based on rbm-lstm neural network for cps in advanced driver assistance system*, ACM Transactions on Cyber-Physical Systems, 4 (2020), pp. 1–17.
- [22] W. XING AND Y. BEI, *Medical health big data classification based on knn classification algorithm*, IEEE Access, 8 (2019), pp. 28808–28819.
- [23] M. YILDIRIM, *Artificial intelligence-based solutions for cyber security problems*, in artificial intelligence paradigms for smart cyber-physical systems, IGI Global, 2021, pp. 68–86.
- [24] L. ZHENG, O. WANG, S. HAO, C. YE, M. LIU, M. XIA, A. N. SABO, L. MARKOVIC, F. STEARNS, L. KANOV, ET AL., *Development of an early-warning system for high-risk patients for suicide attempt using deep learning and electronic health records*, Translational psychiatry, 10 (2020), p. 72.

*Edited by:* Sathishkumar V E

*Special issue on:* Scalability and Sustainability in Distributed Sensor Networks

*Received:* Aug 26, 2023

*Accepted:* Oct 29, 2023



## OPTIMIZATION STUDY OF GRID ACCESS FOR WIND POWER SYSTEM CONSIDERING ENERGY STORAGE

XUBO LE\*, QIUHUA CHEN†, MINJIE ZHU‡, YUCHENG GAO§, BINGYE ZHANG¶, AND AISIKAER||

**Abstract.** This study gives an optimized study with details discussion of the access of wind power grid systems and the energy storage that is high in demand in recent days. The wind power system is a renewable energy resource that can help to meet the need or crisis of energy related to the fuel resources that are being increased in recent times. This is also helpful where energy access related to the power supply is difficult. This may also help the places where the shortage of energy and power supply is faced. There are some issues that are faced at different times that are related to the maintenance and handling of the power grid of the wind turbines. The technical process of handling the machines and the one-time cost of investment on it at the first time is also difficult. There are some benefits of the wind power grid with high energy storage capacity that may help to fulfill the demand for energy that is the main issue of the total system of power supply nowadays. The issues should be mitigated with the help of expert and at the coastal area where there is plenty of continuous flow of wind may be helpful with supply of power supply with wind. These advanced systems hold the potential to mitigate the pervasive energy demand issues plaguing contemporary power supply systems. By expertly addressing these challenges and strategically locating wind power grids, especially in coastal areas with consistent wind flow, the dependable supply of electrical energy can be significantly enhanced, thereby offering an effective solution to the prevailing energy supply challenges of our time.

**Key words:** Power, grid, system, energy, storage, issues, challenges, capacity, renewable, fuel, resources

**1. Introduction.** The intermittency of power generation with the help of wind can be the cause of some challenges for normal operation schedules and in different energy states. The power plants that are related to the Pump Storage are made with wind power plants. It can be helpful for the improvement of the resilience power system and the access to large-scale source distribution has been helpful to maintain a stable operation of the energy system with the help of grid systems.

On the other hand, it is also an important thing to make energy storage units and increments of the consumption process of renewable energy resources and its stable operation with the help of the grid system and wind power system is one of them. The energy storage process has a great advantage that is helpful for balanced load, maintenance of the frequency, and the stability of the grid for voltage along with the buffering of the grid.

This study may help to figure out the energy storage space with consideration of the static security that can be effective for renewable energy resources like wind with great safety with a reduced cost of investment in grid systems.

The grid scale storage plays a very important or vital role in the Net Zero Emission by the year 2040 scenario. This storage system also providing the crucial system and the services that are ranging from the operating of the reserves and the balancing that are short term. It also ranges from the ancillary services and the stability of the grid as well as the investment connected to deferment in the transmission and the distribution of lines. This deferment investment is used for the long – term storage of energy as well as the restoring of the operations that are represented by the grid and following the blackout.

In an era marked by the urgency to transition towards sustainable and renewable energy sources, wind power systems have emerged as a pivotal solution to meet the ever-growing global demand for electricity. The

---

\*State Grid Taizhou Power Supply Company, China

†Goldwind Science & Technology Co., Ltd, Beijing, China, 100176

‡State Grid Taizhou Power Supply Company, China

§Goldwind Science & Technology Co., Ltd, Beijing, China,100176 ([yuchenggaoreess@outlook.com](mailto:yuchenggaoreess@outlook.com))

¶State Grid Taizhou Power Supply Company, China

||Goldwind Science & Technology Co., Ltd, Beijing,China,100176

integration of wind energy into our power grids offers a cleaner, more environmentally responsible alternative to conventional fossil fuel-based electricity generation. However, as wind power gains prominence in the energy landscape, there arises an increasing need for efficient and reliable grid access and energy storage solutions. Wind power systems, powered by the kinetic energy of moving air, offer an abundant and renewable source of electricity. These systems have the potential to significantly reduce greenhouse gas emissions and contribute to the reduction of our carbon footprint, thus playing a pivotal role in mitigating climate change. However, the inherent intermittency and variability of wind resources pose challenges in maintaining a stable and reliable power supply. This is where the synergy of wind power systems and energy storage becomes indispensable.

This research seeks to unravel the complexities and nuances of grid access for wind power systems while considering energy storage as a pivotal component. By investigating the technical, economic, and environmental aspects of this integration, we aim to provide valuable insights into the development of sustainable and reliable energy solutions. As we collectively strive to reduce our reliance on fossil fuels and transition to greener energy sources, the fusion of wind power and energy storage holds the promise of a brighter and more sustainable energy future.

## 2. Objectives.

1. To gain a comprehensive understanding of the fundamental principles and intricacies involved in enabling wind power systems to seamlessly connect with the grid while considering the integration of energy storage solutions.
2. To assess the implications and influence of integrating wind power systems with the electrical grid, particularly in terms of optimizing and enhancing energy storage capabilities.
3. To investigate the challenges those are faced at the time of making grid access for the wind power system
4. To get the solutions by mitigating the issues faced in the making of Grid access of wind power systems in energy storage

**3. Methodology.** A data set with the help of analytical inspection in the form of composition of the literature review section. Here, the application of the secondary data set has been verified with different statistical information and articles that are related to wind power and grid access. The secondary data are authentic and collected from Google Scholar and from different news articles. The secondary qualitative data collection method has been followed and interpretivism philosophy has been used. The deductive approach of the research is used for the present study. The deductive approach helps to connect the theoretical aspect and the scientific collection of data and get a relation between them [18].

Accessing the grid for wind power systems is a complex process that requires coordination with multiple stakeholders, from regulatory authorities to utility companies. It's essential to approach the integration with careful planning and adherence to technical and regulatory requirements for a successful and sustainable wind power project.

**4. Accessing the Grid for Wind Power Systems with the Energy Consideration System.** Grid integration with the help of wind energy is a collection of all the activities that are related to the connection of WPPs with the grid. In the first stage, the planning along with the activities before the GOP is done [9]. The physical connection encompassing the activities that are occurring with the firm of wind with the grid is made. The last and final stage is the operation of systems, which are helpful for the activities occurring after the connection of the grid [7].

Wind energy is incorporated with a unit commitment process and with the help of an economic dispatching process [24]. Forecasting is required for the prediction of hourly variability of wind and that is helpful for energy production. The forecasting lead time has to be reduced as it may be helpful for using the grid in time.

Pumped storage hydropower technology stands as the most widely adopted energy storage method, with significant potential for further expansion across various regions [12]. Batteries, on the other hand, have demonstrated remarkable scalability and are integral to grid-scale energy storage solutions. Over recent years, battery technologies in conjunction with grid-scale storage systems have consistently exhibited substantial market growth and development [7]. In addition to these, alternative storage technologies, such as gravity-based storage and compressed air energy storage, play a comparatively smaller role within the contemporary

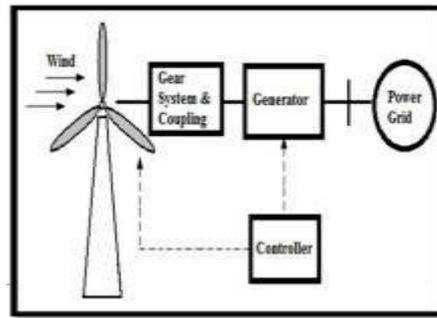


Fig. 4.1: Diagram of wind turbine system

landscape of energy storage systems.

Moreover, the term hydrogen that is detailed separately or differentially is a developing or the growing technology [8]. It is demonstrated as the growing technology that has the additional potential for the seasonal storage related to the renewable energy [13]. While the approaches are still made or projected the development and the growth in the system of the storage capacity or possibility of the grid-scale are not presently on the track in connection with the net zero emissions that need a great hard work and efforts.

**4.1. Impact of Wind Power Plant and Grid System for Energy Storage.** Integration of new power plants can create an impact on the grids that exist and on the incumbent generators. The impact of the multiple time of scale can be described as an hour to hour, day to day, and seconds to second [4].

The impacts of the wind power on the in the system of the grid is mainly depends on the large expansion of the level of the penetration related to the wind energy, the generation mixture of the electricity in the system of the grid-scale, and the size and accuracy of the grid. The expansion of the penetration of the wind-energy at the moderate or the lower level is the matter of the capital investment or the cost.

These issues of the cost are highlighted by numerous regional and national integration studies. This integration of the capital and cost that are highlighted are modest fairly. The low or the moderate penetration levels of the wind energy in the power system operation of the power system are hardly impacted. In recent years 2022, the supply of the power of the wind is less than that of 6 per cent demonstrated by the overall demand of electricity.

There are various establishments of the control procedures and the system reserves that are available for the dealing with variety of the supply and demand that are moderate in dealing with additional variability the lower penetration of the wind energy. This lower penetration of the wind energy levels up an around 30 per cent that are describes depending on the nature of the significant system. In order to get the higher penetration in the level of the wind energy that represents some changes or transformation in the system and the procedures of operations that are required for the accommodation for the further penetration of the winds.

The impacts of the power of the wind energy on the grid power system could be categorized in the long and the short term affects. The long term effects are mainly related to contribution of the wind power that is provided to the adequacy and its capability in order to meet reliably with the problematic situations. Therefore the effects that are short-term are usually caused by the balancing of the system at the time operational scale.

Locally, the wind power plants mostly interact with the voltage of the grid, similarly to any other stations that are related to the power. This context includes the quality of the power, the state steady voltage deviations and the control of the voltage at vor the near the sites of the wind farms and this all must be taken into considerations. The power of the wind provides the control of the voltage and the active power or frequency control. The wind power plants might reduce the distribution and the transmissions losses when applied through as generations that are embedded.

1. Wind power plants can create an impact on the voltage level and the flow of power in different networks. This may be beneficial for the system that is near the load centers and for the low penetration level areas [26].

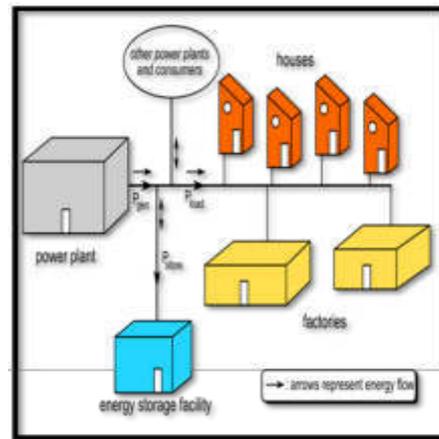


Fig. 4.2: Grid energy storage system

2. Wind power plants may help at the time of low voltage time which can help to improve the up gradation of the transmission and distribution features of current. Connecting with the remote area and sometimes, the pressure of the area that is high in demand of electricity and voltages may be helpful with the help of a wind grid system [23].
3. Wind power has a requirement of measuring and regulating the control and technology generation with a great penetration level in the local area network.
4. If there is the absence of sufficient current and power exchange between the regions and countries that may be helpful with the supply of wind generation and supply [20].
5. Wind power plays an important role in maintenance and stability that contributes to helping the total system with the security of the voltage supply
6. The power plants of the wind can influence the level of the frequency and distribution of the energy in the all the networks. It also supports and benefits the system In the voltage at the time of faults that during the low voltage. This wind plants the reactive power that are able to control all the systems that are installed at the end of the long lines which are radical. This radical lines benefits the system since they support the quality of the voltage as well as the part of the grid.
7. The power of the winds plays a very crucial and vital role in the management of the accuracy and efficiency of the system of the grid as well as the production of the power and storage of the power. It also impacts the security or the safety of the supply of the power in the system of the electricity.

**4.2. Challenges faced by the Wind Integration and making Grid Access for the Wind Power System.** The main primary challenge that is associated with the wind firm is the intermittent nature of winds at different times. Conventional power plants are made and they can give a constant source of voltage that is estimated for distribution [16]. Wind first generated the voltage as per the availability of winds which is highly variable. Therefore, the power generation is uncertain at different times. The quality of the power and voltage is also uncertain as it is dependent on the flow and speed of the winds [25]. The stability of the voltage and the need for it at different times is stable but the wind is not stable the whole time.

Hence, the voltage stability within the wind grid system exhibits fluctuations. Additionally, the construction and installation costs of power plants are notably high, demanding extensive space [2]. Investing such significant capital in a domain characterized by unpredictable wind patterns and power output is not always economically feasible. Furthermore, the grid-related equipment and technologies necessitate expert handling due to their complexity.

There is a lack of technological skills among people to handle the wind grid [17]. The lack of constant current and voltage supply, uncertainty of the amount of the voltage, and the issue those are being faced at the time of making the voltage switch [13]. With the help of wind creates issues to store the current and using it

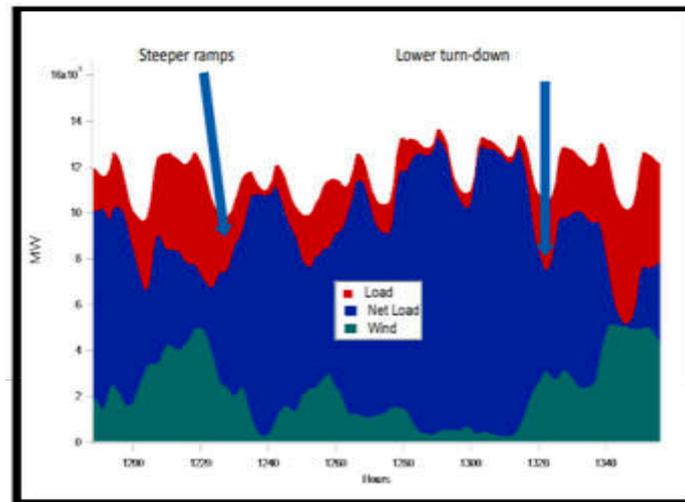


Fig. 5.1: Wind energy has a requirement of flexibility from the generators that are remaining

with the help of renewable energy resources this crisis time of merely. Therefore these challenges are affecting the wind grid power in a great way [5].

**4.3. Solutions for mitigating the Issues faced in making of Grid Access of Wind Power Systems in Energy Storage.** The industry of electricity is going under processing changes worldwide with the help of implementation of the new generation, distribution, and transmission. With the help of consideration, national and regional markets should be the main target for making growth in renewable energy sources [15].

Technologies of new generation on a wider scale can help to mitigate the significant challenges that are faced in the grid system of wind power. In different regions, wind integration issues may be faced over the time of installation of the wind power mills. Sometimes there should be changed the fuel box of the electricity generation process system should be changed in the integrations [14]. There should be invested different technical professionals who have the knowledge about the total process of power grid handling and access in a great way. Training should be provided to the people who are willing to access this system of renewable energy resources [6]. To cut the cost of installation, a huge amount should be given subsidies from the Government to encourage people to use this grid and access it. The awareness of renewable energy resources should be helpful among the people to get interstate and to use it at different times that are useful for power consumption [22].

## 5. Results.

**5.1. Energy storage for the application of grid.** Energy storage can be the asset that may be one of the most valuable assets of the grid system [19]. This can be helpful for the provision of the management of a load of power, quality of power, and uninterrupted power supply that can help to increase the efficiency of the demand of power supply and to get a more sustainable energy system [21].

There are different types of energy storage systems that are suitable for grid-scale applications. Locations of different places are facing the issues. On the other hand, the capacitors used for wind turbines are most effective for conservation of the power and energy [1].

The cost imposed in the total system for renewable energy resources in the market and balancing it with the help of increments of concern of policymakers. The techniques of use and analysis should be assessed in the early stage of the development of the grid power-making stage [11]. The type of approval of the total certification system can also help to mitigate the barriers that are hampering the total process. Storage capacitors should be high technologies that may help to conserve the charges and power [10]. This may help to enhance the energy storage increment that should be used in the time of shortage of power supply or voltage shortage.

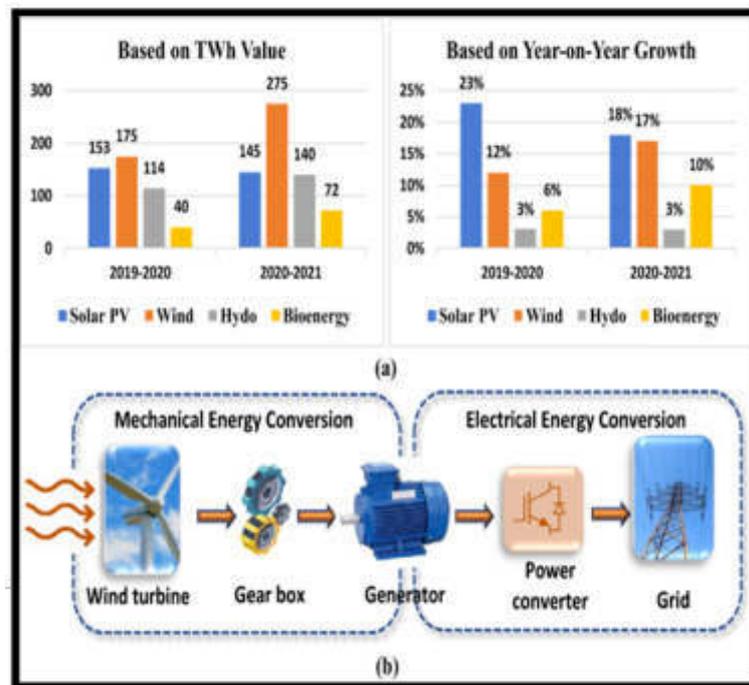


Fig. 5.2: Recent trend of wind energy system

**5.2. Benefits of wind power energy through grid access.** The presence of solar power and an electric grid can create an impact on the natural gas and fire plants by modifying the level of output of variable generation. Fossil fuel generators are non-renewable energy resources that may be finished in the near future [25]. Wind power energy is a clean resource of renewable energy that may be used in the future without fear of being finished.

With the help of wind turbines, the energy is harnessed using the mechanical and technical way with the help of spin generators for creating electricity [3].

Wind energy is an abundant and inexhaustible resource, and it does not pollute the air by burning any fuel. In the United States, the largest energy resource of renewable energy is the supply of wind energy continuously [8]. It helps to become less reliable on renewable energy resources. Wind energy has a great help of 329 million metric tons of carbon emission on an annual basis which is equivalent to 71 million of car emissions that can be the cause of acid rains and greenhouse gasses [12].

As per the basis of lands and regional basis, wind energy and its cost are competitively low in comparison with the other resources. It may be a great cost-saving technical process in the future to get more energy and power and voltage [6]. Wind energy can be implemented in the village or remote areas also as the coastal area may be helpful with the flow of wind. Island communities are other places where the supply of fuel energy is not easily available, but the resources of the wind are plenty and this may be helpful for the place to get the energy supply [4].

The power plants of the wind might need the additional development or growth in the distribution and the transmission of the infrastructure that are related to the grid system. In relation to the case when any of the power plants are connected to the system of the grid.

In order to show connection with the high resources that are remote that includes the off-shore or the air or the wind that are very large in the regions that are remote as well as the centers of the load. The new radical's lines that are need to be constructed for the building of the pipe lines for the gas and oil. This building of the pipeline will maximizing the effects and shows the smooth impacts on the geographically distribution of

the winds. The supply of the power of the wind required many techniques and approaches for regulating and maintaining control similarly any other technology that are related to the generations. This management is necessary for the reduction of the challenges and the issues also increase the power of the firm and include the additional transmission of the energy and power. The absence of the sufficient and the intelligent and the well maintained and the management of the power of the supply and wind and exchange of the power between the countries and the regions are also done through the grid system of the power.

The grid system of the power is describes as the combination or the mixtures of the non-management system that usually demands the production of the of the power that might results in the maintain the efficiency and the frequency of the power supply of the winds where they are constrained.

**6. Conclusion.** In conclusion, wind energy grids represent a crucial component of the modern power storage system, offering the ability to capture and store energy for future use, thereby reducing our reliance on finite fossil fuel resources. Through a comprehensive examination of data collected at various intervals, it becomes evident that the application of energy storage, facilitated by wind energy grids, plays a pivotal role in sustaining a continuous and reliable power supply in regions where traditional sources may fall short. This integration of power grid systems and energy storage, driven by wind energy, underscores its significance as a catalyst for the advancement of innovative technologies. These technologies, in turn, are poised to meet the surging demand for efficient and sustainable power supply systems, ushering in a brighter and more environmentally responsible energy future.

#### REFERENCES

- [1] L. AL-GHUSSAIN, R. SAMU, O. TAYLAN, AND M. FAHRIOGLU, *Sizing renewable energy systems with energy storage systems in microgrids for maximum cost-efficient utilization of renewable energy resources*, Sustainable Cities and Society, 55 (2020), p. 102059.
- [2] A. ATIF AND M. KHALID, *Savitzky-golay filtering for solar power smoothing and ramp rate reduction based on controlled battery energy storage*, IEEE Access, 8 (2020), pp. 33806–33817.
- [3] A. AZIZVAHED, R. KARANDEH, V. CECCHI, E. NADERI, L. LI, AND J. ZHANG, *Multi-area dynamic economic dispatch considering water consumption minimization, wind generation, and energy storage system*, in 2020 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT), IEEE, 2020, pp. 1–5.
- [4] L. BAGHERZADEH, H. SHAHINZADEH, H. SHAYEGHI, AND G. B. GHAREHPETIAN, *A short-term energy management of micro-grids considering renewable energy resources, micro-compressed air energy storage and drps*, International Journal of Renewable Energy Research, 9 (2019), pp. 1712–1723.
- [5] A. BLAKERS, M. STOCKS, B. LU, AND C. CHENG, *A review of pumped hydro energy storage*, Progress in Energy, 3 (2021), p. 022003.
- [6] P. DENHOLM, J. NUNEMAKER, P. GAGNON, AND W. COLE, *The potential for battery energy storage to provide peaking capacity in the united states*, Renewable Energy, 151 (2020), pp. 1269–1277.
- [7] J. A. DOWLING, K. Z. RINALDI, T. H. RUGGLES, S. J. DAVIS, M. YUAN, F. TONG, N. S. LEWIS, AND K. CALDEIRA, *Role of long-duration energy storage in variable renewable electricity systems*, Joule, 4 (2020), pp. 1907–1928.
- [8] D. ENESCU, G. CHICCO, R. PORUMB, AND G. SERITAN, *Thermal energy storage for grid applications: Current status and emerging trends*, Energies, 13 (2020), p. 340.
- [9] A. M. FOLEY, N. MCILWAINE, D. J. MORROW, B. P. HAYES, M. A. ZEHIR, L. MEHIGAN, B. PAPARI, C. S. EDRINGTON, M. BARAN, ET AL., *A critical evaluation of grid stability and codes, energy storage and smart loads in power systems with wind generation*, Energy, 205 (2020), p. 117671.
- [10] M. I. HLAL, V. K. RAMACHANDARAMURTHYA, S. PADMANABAN, H. R. KABOLI, A. POURYEKTA, T. ABDULLAH, AND T. AB RASHID, *Nsga-ii and mopso based optimization for sizing of hybrid pv/wind/battery energy storage system*, International Journal of Power Electronics and Drive Systems, 10 (2019), pp. 463–478.
- [11] A. KADRI, H. MARZOUGUI, A. AOUITI, AND F. BACHA, *Energy management and control strategy for a dfig wind turbine/fuel cell hybrid system with super capacitor storage system*, Energy, 192 (2020), p. 116518.
- [12] H. KHALOIE, A. ABDOLLAHI, M. SHAFIE-KHAH, A. ANVARI-MOGHADDAM, S. NOJAVAN, P. SIANO, AND J. P. CATALÃO, *Coordinated wind-thermal-energy storage offering strategy in energy and spinning reserve markets using a multi-stage model*, Applied Energy, 259 (2020), p. 114168.
- [13] F. A. KHAN, N. PAL, S. H. SAEED, AND A. YADAV, *Modelling and techno-economic analysis of standalone spv/wind hybrid renewable energy system with lead-acid battery technology for rural applications*, Journal of Energy Storage, 55 (2022), p. 105742.
- [14] X. LI AND A. PALAZZOLO, *A review of flywheel energy storage systems: state of the art and opportunities*, Journal of Energy Storage, 46 (2022), p. 103576.
- [15] X. LI AND S. WANG, *Energy management and operational control methods for grid battery energy storage systems*, CSEE Journal of Power and Energy Systems, 7 (2019), pp. 1026–1040.

- [16] M. MAHMOUD, M. RAMADAN, A.-G. OLABI, K. PULLEN, AND S. NAHER, *A review of mechanical energy storage systems combined with wind and solar applications*, Energy Conversion and Management, 210 (2020), p. 112670.
- [17] P. MUKHERJEE AND V. RAO, *Superconducting magnetic energy storage for stabilizing grid integrated with wind power generation systems*, Journal of Modern Power Systems and Clean Energy, 7 (2019), pp. 400–411.
- [18] E. OH AND H. WANG, *Reinforcement-learning-based energy storage system operation strategies to manage wind power forecast uncertainty*, IEEE Access, 8 (2020), pp. 20965–20976.
- [19] D. G. RANGANATHAN ET AL., *Energy storage capacity expansion of microgrids for a long-term*, Journal of Electrical Engineering and Automation, 3 (2021), pp. 55–64.
- [20] U. T. SALMAN, F. S. AL-ISMAIL, AND M. KHALID, *Optimal sizing of battery energy storage for grid-connected and isolated wind-penetrated microgrid*, IEEE Access, 8 (2020), pp. 91129–91138.
- [21] Z. SONG, S. FENG, L. ZHANG, Z. HU, X. HU, AND R. YAO, *Economy analysis of second-life battery in wind power systems considering battery degradation in dynamic processes: Real case scenarios*, Applied Energy, 251 (2019), p. 113411.
- [22] Y. SUN, Z. ZHAO, M. YANG, D. JIA, W. PEI, AND B. XU, *Overview of energy storage in renewable energy power fluctuation mitigation*, CSEE Journal of Power and Energy Systems, 6 (2019), pp. 160–173.
- [23] J. TEH AND C.-M. LAI, *Reliability impacts of the dynamic thermal rating and battery energy storage systems on wind-integrated power networks*, Sustainable Energy, Grids and Networks, 20 (2019), p. 100268.
- [24] Y. TENG, Z. WANG, Y. LI, Q. MA, Q. HUI, AND S. LI, *Multi-energy storage system model based on electricity heat and hydrogen coordinated optimization for power grid flexibility*, CSEE Journal of Power and Energy Systems, 5 (2019), pp. 266–274.
- [25] C. WANG, Z. ZHANG, O. ABEDINIA, AND S. G. FARKOUSH, *Modeling and analysis of a microgrid considering the uncertainty in renewable energy resources, energy storage systems and demand management in electrical retail market*, Journal of Energy Storage, 33 (2021), p. 102111.
- [26] Y. ZENG, C. LI, AND H. WANG, *Scenario-set-based economic dispatch of power system with wind power and energy storage system*, IEEE Access, 8 (2020), pp. 109105–109119.

*Edited by:* Sathishkumar V E

*Special issue on:* Scalability and Sustainability in Distributed Sensor Networks

*Received:* Aug 26, 2023

*Accepted:* Oct 29, 2023



## EVALUATION OF MONITORING TECHNOLOGIES AND METHODS FOR MICRO PLASTICS IN WATER AS NOVEL POLLUTANTS: THE EXPLORATION OF ACCURATE QUANTITATIVE ANALYSIS AND EFFICIENT SCREENING

KE HU, DONGDONG LI, XIAOLEI CUI, DONGHUA HU, JUNLIANG CHEN, SHAOPENG ZHUAN, HAO CHANG, YAPING ZHANG, TINGTING AN, AND JUQIN ZHANG\*

**Abstract.** Micro plastics have recently emerged as a major biohazard that has a considerable impact on the environment. Moreover, of the detrimental capabilities of micro plastics, a hope of controlling efforts of micro plastics has been in the headlines. microplastics have gained notoriety due to their adverse effects on the environment and wildlife. Controlling these minuscule yet harmful particles requires effective monitoring, detection, and management strategies. This analysis delves into the diverse techniques and technologies available for tracking and mitigating microplastic pollution. Therefore, the following analysis has aimed at analysing the monitoring technologies and methods for micro plastics. Additionally, the monitoring methods are observed along with the advantages and disadvantages. For the development of the analysis, a secondary qualitative method was used in the process. Additionally, the graphical representation of the efforts for controlling the novel pollutant is analysed along with relevant problems. Hence, a coherent discussion is presented in the following analysis. This research contributes to the broader understanding of microplastic pollution and its monitoring while underlining the need for enhanced control measures. It provides a valuable resource for policymakers, environmentalists, and researchers working toward a cleaner, more sustainable environment. As microplastics continue to infiltrate ecosystems worldwide, comprehensive monitoring and control efforts are of paramount importance.

**Key words:** Micro plastic, Ocean pollution, Screening, Sources

**1. Introduction.** Microplastics, minuscule plastic particles less than 5 millimeters in size, have become a global environmental concern, particularly when they infiltrate aquatic ecosystems. These tiny plastic fragments have emerged as significant pollutants in our water bodies, presenting a range of ecological and human health challenges. The proliferation of microplastics can be attributed to the breakdown of larger plastic items, like bottles and packaging, and the widespread use of microbeads in personal care products. Additionally, microplastics can originate from the wear and tear of synthetic textiles, tire abrasion, and the degradation of marine debris

Due to the properties of micro plastic, it is hard to detect the presence of the same in the environment. Therefore, development in the methods of detecting micro plastic plays a major role in environmental protection. It was estimated that around 92% of plastic particles are present in the ocean as a major pollutant [1]. Additionally, it was noted that such a heavy amount of plastic is harmful to ocean wildlife. Thus, the following analysis has aimed at evolving the monitoring technologies used for micro plastic detection [2]. Further, advantages and disadvantages are presented in the following analysis in order to determine the impact of the methods.

In figure 1.1, micro plastic monitoring comprises of various waste particles that have been reported as harmful around the globe to understand the role of waste water which is increasing abruptly. In this section, the various array based reflections have been made that are based on microbes and are primarily settled due to 76.9% of the total MP detectors [3]. With the increase in micro plastic pollution, the rivers and seas are suspected to lose their nutritional values soon thereby harmony the lives of the water bodies predominantly. As a result of this, the various harmful factors are addressed to have concomitant outcomes that are found to be difficult in practice.

The detrimental impact of microplastics is multifaceted. They pose a direct threat to marine life as they are often mistaken for food and ingested by a variety of organisms, from zooplankton to fish, eventually entering

---

\*The First Geological Brigade of Hebei Bureau of Geology and Mineral Resources Exploration, HanDan, 056000, China (Corresponding author: [dongdonglires@outlook.com](mailto:dongdonglires@outlook.com))

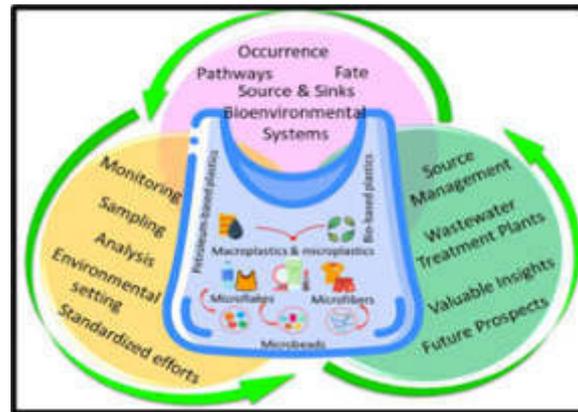


Fig. 1.1: Graphical representation of micro plastic monitoring

the food chain. Beyond the ecological consequences, microplastics are also a matter of human health concern as they can carry harmful chemicals and transfer them to the seafood we consume. It is imperative to comprehend the extent of microplastic pollution in water, monitor its presence, and develop strategies for its mitigation. This paper offers a detailed analysis of microplastic monitoring technologies and methods, shedding light on their advantages and disadvantages, and emphasizing their significance in addressing this modern environmental challenge.

**2. Objectives.** In order to develop the results in an effective manner the following analyses objectives were followed:

- To provide a comprehensive analysis of microplastics and their influence on aquatic ecosystems.
- To examine the methodologies employed in microplastic monitoring.
- To assess the existing limitations in environmental microplastic detection technologies.
- To explore the drawbacks of various monitoring methods and their implications for scalability.

**3. Methodology.** The methodology of analysis looks into the elements that contributed to the development of the empirical analysis. Therefore, it can be contemplated that the methodology of a study is responsible for the systematic development of tangible results [3]. In order to analyse the monitoring methods of microplastics secondary data was considered in the study. Furthermore, a qualitative method of analysis was employed in the following analyses. Collecting secondary information can be beneficial in analysing reliable information [4]. Additionally, with the use of the qualitative method of analysis, it was possible to implement a reliable relation among different factors. Further, advantages and disadvantages were analysed and reliable methods are prescribed based on the relations.

Microplastics are tiny plastic particles or fibers measuring less than 5 millimeters in size. They can come from various sources and can be broadly categorized into two main types: primary microplastics and secondary microplastics.

1. Primary Microplastics:

- (a) Microbeads: These are small plastic particles commonly found in personal care products such as exfoliating scrubs and toothpaste. When washed down the drain, they can enter water bodies, posing a threat to aquatic life.
- (b) Nurdles: Nurdles are pre-production plastic pellets used in manufacturing. Accidental spills during transportation or handling can lead to their entry into water systems.
- (c) Microfibers: Microfibers are tiny threads shed from synthetic textiles when we wash our clothes. They are a major contributor to microplastic pollution in oceans.

2. Secondary Microplastics:

- (a) Fragmented Plastics: Larger plastic items, like bottles and bags, can break down over time into

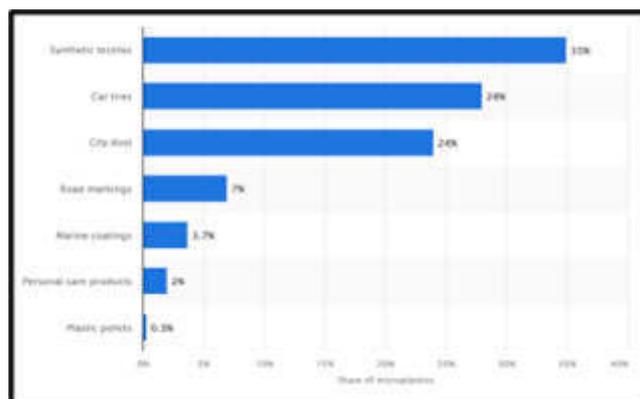


Fig. 3.1: Sources of micro plastic and their share

smaller fragments due to weathering processes. These fragments then become secondary microplastics.

- (b) Tire Wear Particles: As vehicle tires wear down, they release small rubber particles containing microplastics into the environment. These are often washed into water bodies by rain.
- (c) Plastic Film Fragments: Discarded plastic bags and packaging materials can degrade into smaller pieces that find their way into water systems.

The harm caused by microplastics includes ingestion by marine life, with potential health effects extending up the food chain to humans. These particles can disrupt ecosystems, harm wildlife, and introduce toxins into aquatic environments. Addressing microplastic pollution is crucial to protect both the environment and human health.

**3.1. Micro plates and their impact on environments .** Micro plastics have emerged as a major micro pollutant over the years. Primarily micro plastics are small plastic particles that are not destroyed and cause pollution in the environment. Micro plastics range in size from less than 5 millimetres to as small as a few nanometres [5]. They are a kind of plastic waste that may be discovered in a variety of environmental contexts, such as soil, water, and even the air [6]. The disintegration of bigger plastic objects, the breakdown of synthetic fabrics, and the discharge of micro beads from personal care products are just a few of the potential sources of micro plastics. It was noted that around 92% of micro plastics are found on the ocean surface [7]. Therefore, micro plastic can be contemplated as one of the major pollutants of aquatic wildlife.

In the above graph 3.1, the different sources of micro plastics are illustrated along with their share of ocean micro plastic. It can be seen that synthetic textile is one of the major pollutants that have a 35% share. In addition, was such as car tires and dust are one of the major contributors to the increase in micro plastic particles in the ocean. Additionally, personal care products have a 2% share of increasing the plastic in ocean beds [8].

It was noted that there are two kinds of micro plastic that are responsible for the pollution.

**The primary source of micro plastic:** Primary micro plastics can be defined as plastic particles that are purposefully produced in a small size, such as plastic pellets used in industrial operations or micro beads used in cosmetics and cleaning goods 9.

**Secondary sources of micro plastics:** Secondary micro plastic are plastic fragments that are left over when bigger plastic objects deteriorate and break apart over time [9]. For instance, environmental elements like sunshine, wind, and wave action can cause the breakdown of plastic bottles, bags, and other plastic items, resulting in the production of smaller particles.

**3.2. Methods for monitoring micro plastic.** Due to the growing identification of micro plastics as new pollutants with potentially adverse effects on the environment and human health, the assessment of monitoring technologies and methodologies for micro plastics in water is of utmost relevance. Understanding the distri-

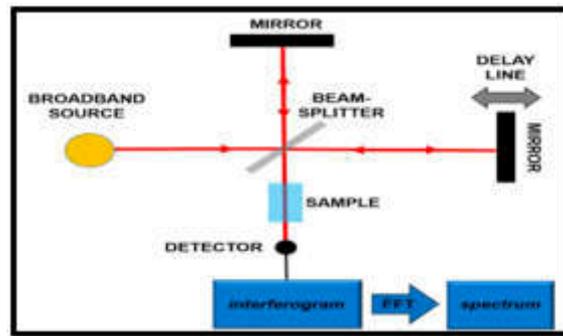


Fig. 3.2: Fourier Transform Infrared Spectroscopy (FTIR)

bution, concentration, and origins of micro plastics in aquatic habitats requires accurate quantitative analysis and effective screening. Following are some of the methods used for micro plastic monitoring:

**“Fourier Transform Infrared Spectroscopy (FTIR)”**: Based on their molecular vibrations, FTIR can distinguish between various polymer types in micro plastics [10]. It may be used in conjunction with other quantification techniques and is helpful for qualitative analysis.

**Raman Spectroscopy**: Raman spectroscopy has the benefit of being non-destructive and, similar to FTIR, can distinguish different kinds of polymers [11]. Both qualitative and quantitative analyses can benefit from Raman spectroscopy [3].

**Method of optical microscopy**: Micro plastics can be visually identified and quantified using optical microscopy, including stereomicroscope and microscopy employing dyes like Nile Red [12]. Although time-consuming and somewhat limited in its ability to identify microscopic particles, it is reliable [13].

**Method of Sediment sampling**: Sediment is a type of testing that is helpful in detecting the excessive amount of plastics under the sea level. It is performed by extracting the plastics and products of plastics by the potential sources with the help of commercial ships and aquaculture activities [16]. Through sediment sampling, the depth of the sea bodies can be estimated and the microscope reviver rate is expected to raise with the different pre treatment methods. The extraction of micro plastics may include both the steps in which natural organic materials are degraded resulting in rinsing off the saturated solution segments [18].

**Organic degradation**: Degradation can be possible when microorganisms such as the fungi and bacteria with the help of enzymatic actions convert themselves into the metabolic products such as carbon dioxide, methane and water respectively [19]. It includes the initial attachments off microbial and micro plastics with the assistance of the bio degradable energies that are the outcome off microbes.

**Dye natural fibres**: The influence of common textile based on cotton fabrics are assessed to be durable that generates nearly 14,000 particles of cotton fabric. The microbes are formed under water by cross linking with the repellent treatment to contribute into loosing the microfiber shedding in the global marine environment [21]. However, a mismatched plastic waste ends up entering into the water bodies causing miserable harm to the water bodies. The practice of bio degradation is worldwide and it helps in dumping waste and litters into smaller fragments to replenish the micro plastics as well as to measure the density of the water.

**Sieve supernatants**: The separation of harmful particles from the water bodies has made new techniques and efficacies to separate micro plastics from the matrix thick and rich mass of obligations. The abundance of micro plastics is pertinent to be limiting the associated hazards and thereby increasing the ample of resources to factories the reliability of the obtained results in a significant manner [23]. It is relevant to the micro plastics that can be compared relatively by isolating the micro plastics from the mass of environmental samples in order to get the appropriate results.

**Auto sampler**: Auto sampler is use for the collection of samples that are precise in the artificial practices of collecting the samples of the dust particles to gain utmost productivity. It is an automated machine that is associated with the gas chromatographers that revolves in the horizontal position. Auto sampler liquids

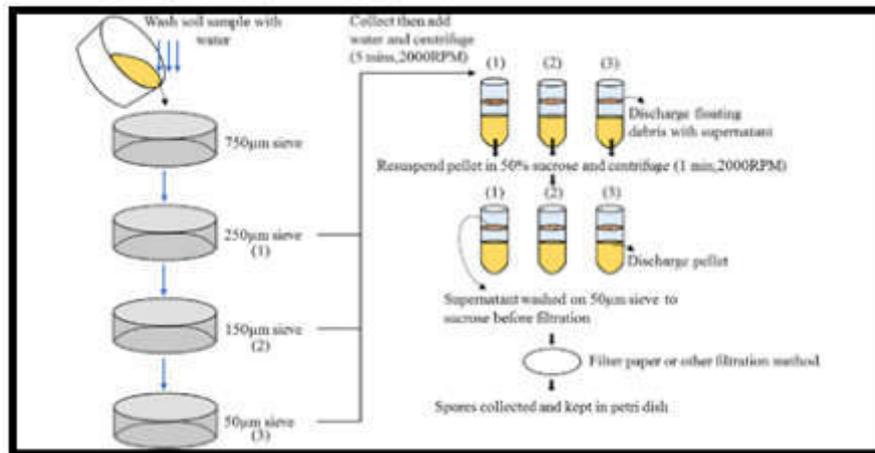


Fig. 3.3: Sieve supernatants of micro particles

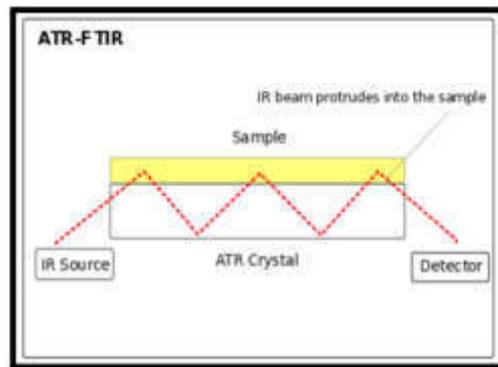


Fig. 3.4: ATR FTIR analysis

comprises of concentric hoardings that enable the pumping syringe to push the bio resources into the water bodies. The sampling apparatus is the titration that determines the carousel auto samplers to inhibit into the potential courses to acknowledge the floating of discharged liquid that are commonly used in the gas chromatography [22].

**ATR FTIR analysis:** Attenuated total reflectance(ATR), is the most widely is sampling method that determines the transformation of water bodies particles to administer the range of samples to form to explore the natural differences to explore the natural habitat that are accustomed to cluster the organic functions to control the water molecules in order to determine the frequency of the water bodies [24]. The mineral resources are observed to articulate the various concomitants in a selective manner.

The above figure 3.4 is illustrates the functional elements of the ATR cycle that reflects the IR beam samples which is protruded to verify the IR sources. The relevant practiced are determined to acknowledge the functional operations in a significant manner.

**3.3. Gaps in the process of Micro plastic monitoring.** During the analysis of past literature, it was noted that there are some specific issues related to micro plastic detection [14]. One of the major factors was related to the cost of the process. It was noted that the cost of FTIR and Raman Spectroscopy was one of the major hindrances to the mass implication of such methods. However, the effectiveness of the methods is

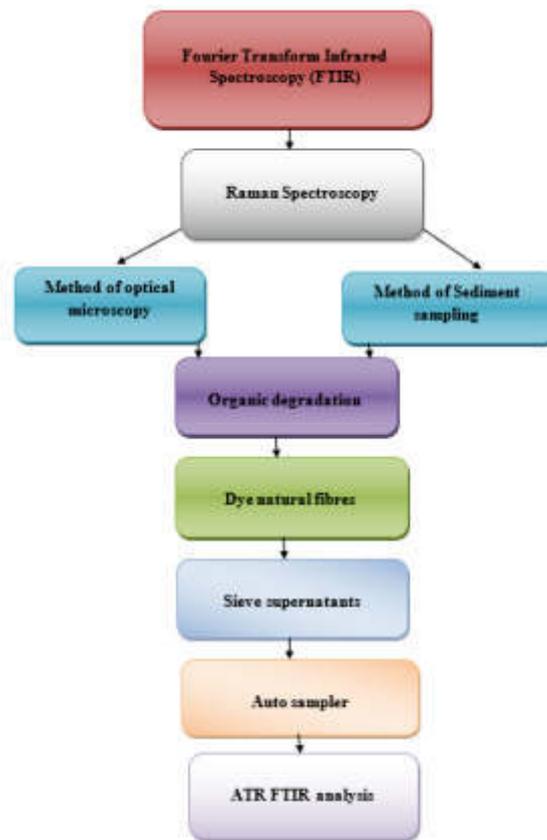


Fig. 3.5: Flow chart analysis of micro plastic

comparatively better than other methods. Additionally, it was determined that such methods are complex in nature [15]. Hence, it can be contemplated that the complexity of the process of major decision-making and data interpretation can be impacted.

**4. Results.** In order to determine the monitoring process of micro plastic a systematic analysis was conducted where different factors were observed. It was noted that micro plastics are one of the major issues that are polluting the ocean beads. Therefore, accurate monitoring is essential in order to improve the condition [16].

In the above illustration 4.1, a predictive analysis is presented that indicates the growth of micro plastic pollution by the year 2050 [17]. It can be seen that with appropriate precautions it is possible to stop the emission rate by the year 2020. However, the above graph further indicated that falling to control the micro plastic can result in a 2.5 million tone of micro plastics in the ocean [18]. Therefore, developing effective monitoring is essential for the process.

From the above graphical 4.2 representation, a detailed understanding of the surface micro plastic can be comprehended. It can be contemplated that there are different levels of the sink for the Micro plastics [19]. Hence, it was found that the monitoring methods need to efficiently identify the different levels of micro plastic identification. Moreover, for such implications, micro plastic need to methods need to be cost-effective [20]. At the time of analysing pat literature, it was noted that there are certain issues related to the methods that are hindering the mass application [21]. For instance, costly processes, complex methods, and interpretation of accurate information are major disadvantages of FTIR methods [22].

However, some of the modern methods are able to counter such issues in an effective manner. For instance, Method of holographic identification where a 3D imaging of the particle is created [23]. In such a manner, a

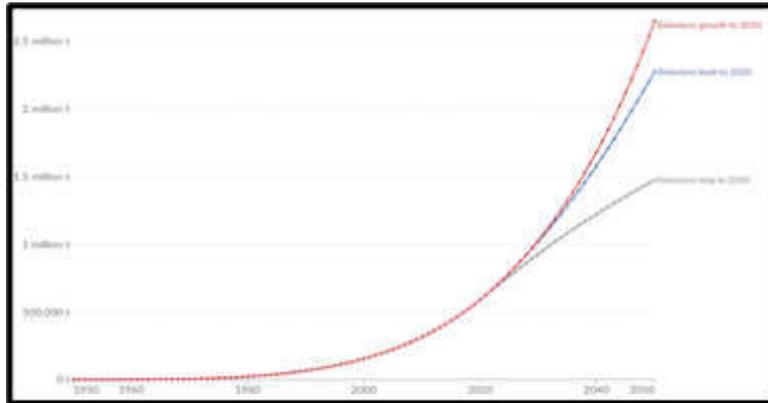


Fig. 4.1: Predictive analysis of the possible directions of micro plastic pollution

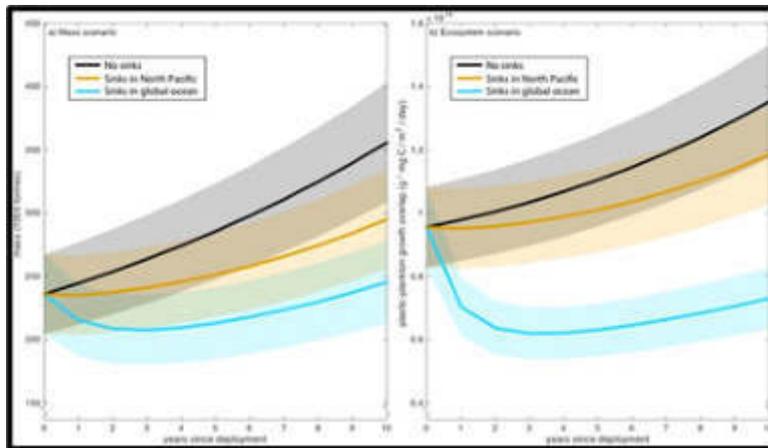


Fig. 4.2: Time history showing (a) the total mass of marine micro plastic present at the ocean’s surface with and without sinks (black line), sinks present in the North Pacific (yellow line), and sinks present close to coasts (blue line).

better prediction of the micro plastic can be conducted. Furthermore, Nano particle methods aid to identify the small sizes of plastic [24]. Therefore, an effective contemplation of the particles can be conducted. Additionally, methods of collecting samples have a major impact on the Identification method of the data [25]. It was found that with appropriate sample collection size, range, and propagation of the particles can be identified [26].

The above figure 4.3 signifies the probability rate of the micro particles that are addressed as detectable to produce logarithmic scales. The popularity of the instrumental conditioning have enforced the various courses of actions to acknowledge the supplementary data sets that can be exposed to the rising commercial possibilities in a significant manner [26]. The maximum water level is measured to be cumulative in nature to measure the uncertainties of micro plastics being exposed into the water bodies. The concentrated part is laid off to factorise the small vesicles forming plasma membrane around the water bodies. The cumulative frequency raises to 1.0 with a concentration of activating platelets to compliment the activating exposure to over activate the homeostasis of the circulating micro particles [25].

The above figure 4.4 illustrates the co exposure of viruses developed from the micro plastics that abnormally increases the mortality rate of the fish and other water animals. The graphical presentation shows that with zero micro particles, the water remains harmless and clear as crystal. The moments it is accompanies with

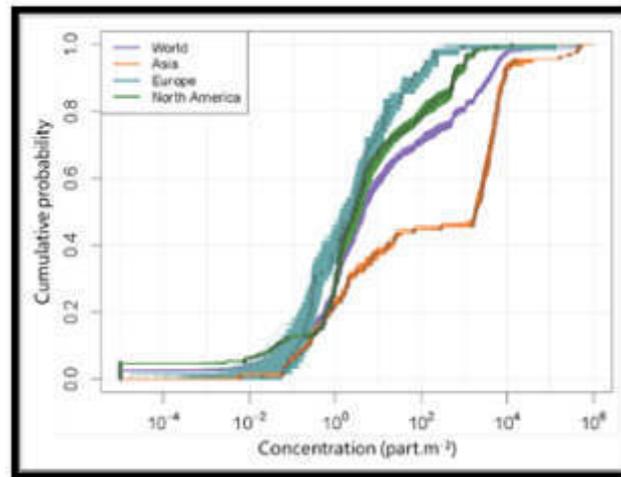


Fig. 4.3: Probability curve of the micro particles

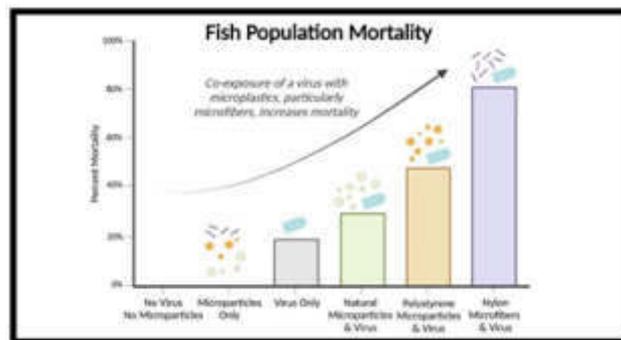


Fig. 4.4: Affectivity rate of the water bodies due to micro particles

virus and nano particles, it decreases the proximity level of the water making it harmful and poisonous for the water bodies [18].

The above figure 4.5 illustrates the global annual plastic dumping into the water bodies and carcinogenic substances. This reduces the fertility of the water thereby making it less effective for the water living bodies. The prevalence of viruses has reduced the quality of the water thereby making it polystyrene in condition [19]. The credibility level reduces and the nylon micro fibres incrses the mortality rate of the water bodies by exposing them to the harmful gases.

**5. Conclusion.** Therefore, the following analysis conducted a secondary analysis on the impact of micro plastics and effective monitoring methods of the same. For instance, it was found that monitoring methods such as FTIR and Raman Spectroscopy are one of the major monitoring methods for micro plastic in ocean beds. Additionally, it was noted that there are different sinking stages of the micro plastic that need to be monitored in an effective manner. Gaps in the process were discussed in an effective manner that sheds light on the disadvantages of current methods. It was noted that cost-effectiveness and accuracy is the prime hindrance to the scalability of such systems. Therefore, a coherent discussion on the use of monitoring technology for micro plastic in ocean beds is presented in the study.

In conclusion, this study underscores the critical importance of microplastic monitoring in addressing the

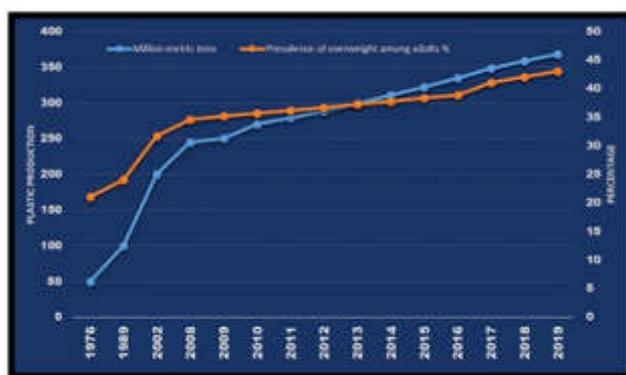


Fig. 4.5: Dumping of plastics in water bodies annually

growing environmental crisis posed by these minuscule yet highly damaging particles. Microplastics have emerged as a significant biohazard, profoundly impacting both aquatic ecosystems and the health of organisms within them. The systematic analysis conducted here has shed light on the various aspects of microplastic monitoring, from understanding their sources and distributions to evaluating the effectiveness of different methods and technologies. The findings have revealed the urgent need for accurate monitoring and mitigation efforts to curb the escalating pollution of water bodies, particularly the oceans.

The predictive analysis presented in this study serves as a stark reminder of the trajectory of microplastic pollution if immediate action is not taken. While the possibility of halting the emission rate by 2020 exists, the consequences of failing to do so are dire, with projections of 2.5 million tons of microplastics flooding our oceans by 2050. It is imperative that we heed these warnings and work towards implementing effective monitoring strategies to control the growing menace of microplastics. Additionally, the results indicate that different methods for monitoring microplastics come with their own advantages and disadvantages.

#### REFERENCES

- [1] R. BEIRAS AND A. M. SCHÖNEMANN, *Currently monitored microplastics pose negligible ecological risk to the global ocean*, *Scientific reports*, 10 (2020), p. 22281.
- [2] M. BIGDELI, A. MOHAMMADIAN, A. PILECHI, AND M. TAHERI, *Lagrangian modeling of marine microplastics fate and transport: The state of the science*, *Journal of Marine Science and Engineering*, 10 (2022), p. 481.
- [3] F. CHAI, K. S. JOHNSON, H. CLAUSTRE, X. XING, Y. WANG, E. BOSS, S. RISER, K. FENNEL, O. SCHOFIELD, AND A. SUTTON, *Monitoring ocean biogeochemistry with autonomous platforms*, *Nature Reviews Earth & Environment*, 1 (2020), pp. 315–326.
- [4] V. H. DA SILVA, F. MURPHY, J. M. AMIGO, C. STEDMON, AND J. STRAND, *Classification and quantification of microplastics (< 100  $\mu\text{m}$ ) using a focal plane array-fourier transform infrared imaging system and machine learning*, *Analytical Chemistry*, 92 (2020), pp. 13724–13733.
- [5] J. GAGO, A. FILGUEIRAS, M. L. PEDROTTI, M. CAETANO, AND J. FRIAS, *Standardised protocol for monitoring microplastics in seawater. deliverable 4.1.*, (2019).
- [6] B. HUFNAGL, M. STIBI, H. MARTIROSYAN, U. WILCZEK, J. N. MÖLLER, M. G. LÖDER, C. LAFORSCH, AND H. LOHNINGER, *Computer-assisted analysis of microplastics in environmental samples based on  $\mu\text{ftir}$  imaging in combination with machine learning*, *Environmental science & technology letters*, 9 (2021), pp. 90–95.
- [7] C. HUNG, N. KLASIOS, X. ZHU, M. SEDLAK, R. SUTTON, AND C. M. ROCHMAN, *Methods matter: methods for sampling microplastic and other anthropogenic particles and their implications for monitoring and ecological risk assessment*, *Integrated environmental assessment and management*, 17 (2021), pp. 282–291.
- [8] J. S. JONES, A. GUÉZOU, S. MEDOR, C. NICKSON, G. SAVAGE, D. ALARCÓN-RUALES, T. S. GALLOWAY, J. P. MUÑOZ-PÉREZ, S. E. NELMS, A. PORTER, ET AL., *Microplastic distribution and composition on two galápagos island beaches, ecuador: Verifying the use of citizen science derived data in long-term monitoring*, *Environmental Pollution*, 311 (2022), p. 120011.
- [9] P. KERSHAW, A. TURRA, F. GALGANI, ET AL., *Guidelines for the monitoring and assessment of plastic litter and microplastics in the ocean.*, (2019).
- [10] S. LECHTHALER, L. HILDEBRANDT, G. STAUCH, AND H. SCHÜTTRUMPF, *Canola oil extraction in conjunction with a plastic free separation unit optimises microplastics monitoring in water and sediment*, *Analytical Methods*, 12 (2020), pp. 5128–5139.

- [11] C. LI, R. BUSQUETS, AND L. C. CAMPOS, *Assessment of microplastics in freshwater systems: A review*, *Science of the Total Environment*, 707 (2020), p. 135578.
- [12] C. MASSARELLI, C. CAMPANALE, AND V. F. URICCHIO, *A handy open-source application based on computer vision and machine learning algorithms to count and classify microplastics*, *Water*, 13 (2021), p. 2104.
- [13] A. MELET, P. TEATINI, G. LE COZANNET, C. JAMET, A. CONVERSI, J. BENVENISTE, AND R. ALMAR, *Earth observations for monitoring marine coastal hazards and their drivers*, *Surveys in Geophysics*, 41 (2020), pp. 1489–1534.
- [14] Y. MICHIDA, S. CHAVANICH, S. CHIBA, M. R. CORDOVA, A. COZSAR CABANAS, F. GLAGANI, P. HAGMANN, H. HINATA, A. ISOBE, P. KERSHAW, ET AL., *Guidelines for harmonizing ocean surface microplastic monitoring methods. version 1.1.*, (2019).
- [15] E. MILLER, M. SEDLAK, D. LIN, C. BOX, C. HOLLEMAN, C. M. ROCHMAN, AND R. SUTTON, *Recommended best practices for collecting, analyzing, and reporting microplastics in environmental media: Lessons learned from comprehensive monitoring of san francisco bay*, *Journal of hazardous materials*, 409 (2021), p. 124770.
- [16] S. PHAN, D. TORREJON, J. FURSETH, E. MEE, AND C. LUSCOMBE, *Exploiting weak supervision to facilitate segmentation, classification, and analysis of microplastics (< 100  $\mu\text{m}$ ) using raman microspectroscopy images*, *Science of The Total Environment*, 886 (2023), p. 163786.
- [17] S. PRIMPKE, A. M. BOOTH, G. GERDTS, A. GOMIERO, T. KÖGEL, A. LUSHER, J. STRAND, B. M. SCHOLZ-BÖTTCHER, F. GALGANI, J. PROVENCHER, ET AL., *Monitoring of microplastic pollution in the arctic: recent developments in polymer identification, quality assurance and control, and data reporting*, *Arctic Science*, 9 (2022), pp. 176–197.
- [18] A. SHAHANAGHI, Y. YANG, AND R. M. BUEHRER, *Stochastic link modeling of static wireless sensor networks over the ocean surface*, *IEEE Transactions on Wireless Communications*, 19 (2020), pp. 4154–4169.
- [19] F. STOCK, C. KOCHLEUS, B. BÄNSCH-BALTRUSCHAT, N. BRENNHOLT, AND G. REIFFERSCHIED, *Sampling techniques and preparation methods for microplastic analyses in the aquatic environment—a review*, *TrAC Trends in Analytical Chemistry*, 113 (2019), pp. 84–92.
- [20] L. W. VAN RAAMSDONK, M. VAN DER ZANDE, A. A. KOELMANS, R. L. HOOGENBOOM, R. J. PETERS, M. J. GROOT, A. A. PEIJNENBURG, AND Y. J. WEESEPOEL, *Current insights into monitoring, bioaccumulation, and potential health effects of microplastics present in the food chain*, *Foods*, 9 (2020), p. 72.
- [21] S. WANG, Q. DONG, L. DUAN, Y. SUN, M. JIAN, J. LI, AND J. DONG, *A fast internal wave detection method based on pcanet for ocean monitoring*, *Journal of Intelligent Systems*, 28 (2019), pp. 103–113.
- [22] S. L. WRIGHT, T. GOUIN, A. A. KOELMANS, AND L. SCHEUERMANN, *Development of screening criteria for microplastic particles in air and atmospheric deposition: critical review and applicability towards assessing human exposure*, *Microplastics and Nanoplastics*, 1 (2021), pp. 1–18.
- [23] J. YANG, J. WEN, Y. WANG, B. JIANG, H. WANG, AND H. SONG, *Fog-based marine environmental information monitoring toward ocean of things*, *IEEE Internet of Things Journal*, 7 (2019), pp. 4238–4247.
- [24] C. ZHANG, L. LIU, L. ZHOU, X. YIN, X. WEI, Y. HU, Y. LIU, S. CHEN, J. WANG, AND Z. L. WANG, *Self-powered sensor for quantifying ocean surface water waves based on triboelectric nanogenerator*, *Acs Nano*, 14 (2020), pp. 7092–7100.
- [25] X. ZHANG, H. ZHANG, K. YU, N. LI, Y. LIU, X. LIU, H. ZHANG, B. YANG, W. WU, J. GAO, ET AL., *Rapid monitoring approach for microplastics using portable pyrolysis-mass spectrometry*, *Analytical chemistry*, 92 (2020), pp. 4656–4662.
- [26] Y. ZHU, C. H. YEUNG, AND E. Y. LAM, *Microplastic pollution monitoring with holographic classification and deep learning*, *Journal of Physics: Photonics*, 3 (2021), p. 024013.

*Edited by:* Sathishkumar V E

*Special issue on:* Scalability and Sustainability in Distributed Sensor Networks

*Received:* Aug 26, 2023

*Accepted:* Oct 28, 2023



## A METHOD FOR SPECIFYING YOGA POSES BASED ON DEEP LEARNING, UTILIZING OPENCV AND MEDIA PIPE TECHNOLOGIES

ANURADHA T\*, N. KRISHNAMOORTHY†, C S PAVAN KUMAR‡, L V NARASIMHA PRASAD §, CHUNDURU  
ANILKUMAR ¶, AND USHA MOORTHY\*||

**Abstract.** Yoga is a years discipline that calls for physical postures, mental focus, and deep breathing. Yoga practice can enhance stamina, power, serenity, flexibility, and well-being. Yoga is currently a well-liked type of exercise worldwide. The foundation of yoga is good posture. Even though yoga offers many health advantages, poor posture can lead to issues including muscle sprains and pains. People have become more interested in working online than in person during the last few years. People who are accustomed to internet life and find it difficult to find the time to visit yoga studios benefit from our strategy. Using the web cameras in our system, the model categorizes the yoga poses, and the image is used as input. However, the media pipe library first skeletonizes that image. Utilizing a variety of deep learning models, the input obtained from the yoga postures is improved to improve the asana. The algorithms like VGG16 (Visual Geometric Group), VGG19, Convo2d, CNN.

**Key words:** Deep Learning, Media Pipe, OpenCV, VGG16, VGG19, Skeletonization.

**1. Introduction.** Yoga is a type of exercise is not just those looking to stay fit and healthy but for those who seek harmony in their lives. An admixture of asanas and postures, yoga is aimed to help with inflexibility, attention, and calmness. From bending and stretching to contemplation, yoga helps to regulate ails, reduce weight, and attack issues like stress and anxiety. Yoga is the ideal drug for those who like to lead healthy lives and make a connection with body, mind, and spirit. Mortal posture assessment is a delicate issue in the control of PC vision. It manages confinement of mortal joints in a picture or videotape to shape a cadaverous depiction. To accordingly fete an existent's posture in a picture is a worrisome errand as it relies upon colorful perspectives, for illustration, scale and thing of the picture, enlightenment variety, foundation mess, dress kinds, environmental factors, and connection of people with the environmental factors. The external gadget with a built-in algorithm that recognizes body corridors and markers as a result in each depth image. The approach is based on arbitrary decision trees that were trained from a sizable database of artificial body part marker images. Several mortal body models and a database of mortal stir prisoner data are used to create the synthetic training images. The diversity and unpredictability of the stir prisoner data reveal the efficiency of the body part recognition system. For instance, the training data set's unbalanced distribution of data in one stir delicacy bloody compared to the other orders will have a poor impact on the recognition performance. It should be carefully designed to collect the training data somewhat from the area of mortal activities. In real life, negotiating this is difficult. Many locomotion's, gestures, and standing behaviors can be easily acquired using marker-grounded stir training and are included in stir databases that are open to the public. However, because of marker occlusion, landing behaviors that need the person to thicken, sit, kneel, or bend are more taxing and do not yield as many data as standing behaviors. A sizable collection of stir data from various stir orders is used to train the recognition system, however only some stir orders may have enough diversity. The foundation for posture evaluation and recognition is the information-gathering of the entire body. In order to collect posture information using colorful detectors, numerous biases have been devised. The most popular detectors are inertial dimension units, wearable entrance markers, picture detectors, and electromyogram detectors (IMUs).

---

\*Department of Information Technology, VR Siddhartha Engineering College, Vijayawada, Andhra Pradesh, India

†Department of Computer Science and Engineering, Kongu Engineering College, Perundurai, Erode, Tamilnadu 638060, India

‡Department of Information Technology, VR Siddhartha Engineering College, Vijayawada, Andhra Pradesh, India.

§Department of Computer Science and Engineering, Institute of Aeronautical Engineering, Hyderabad, Telangana, India.

¶Department of Information Technology, GMR Institute of Technology, Rajam, Srikakulam, Andhra Pradesh, India.

||Department of Computer Science and Engineering, REVA University, Bengaluru, India ([usha.m@reva.edu.in](mailto:usha.m@reva.edu.in))

In general, people do not need to wear redundant equipment while using image detectors, picture other side be easy body pose should not cause image occlusion

The contribution of the paper is summarized below: -

1. A useful system for categorizing yoga positions that can be practiced at home unaccompanied. The general pose discovery model is replaced by the precise and reliable Google Media Pipe library for body key point recognition.
2. To make the system less amenable to different conclusions, the model employs training and testing images with different backgrounds. The system allows for background changes and disturbances, such as exposure to bright light, because the background in each image is the same.
3. It makes use of a sizable dataset made up of pictures of yoga poses for testing, training, and validation.
4. Future research on yoga pose detection will be made possible by comparative analysis of the various models.
5. A precise and low-latency model is provided for real-time settings for practitioners who are housebound.

**2. Related Work.** The 3D CNN architecture's planned structural elements. The 3D CNN armature's suggested structural elements. also shows the developed network's configuration data, including the large count of parameters corresponding to the different layers, kernel sizes, and the overall size of the point maps. As depicted in the mentioned figure, the input to the specially constructed containing 3D CNN as an image cell made up of sixteen frames with a resolution of containing 112 x 112 x 3 pixels that were taken from the video's inputs of yoga poses. Neural network, which is made up of different layers 8 Conv3D layers ,5 MaxPool3D layers and AvgPool3D subcaste, reuses the input picture cell after classifying one of the ten Yoga actions using a SoftMax classifier. Each of the 3 x 3 x 3 kernels used in the convolutional layers corresponds to aimlessly initialized pollutants. During convolution, the pollutants apply a stride of one along the spatial and temporal boundaries with border mode set as valid. For the network to increasingly prioritize discriminating features from the input video of the Yoga positions, the number of pollutants in the convolutional layers is gradually raised from 64 to 512 [12]. Our slice technique produces stir data, which is used to train arbitrary decision trees that automatically annotate input depth photos. The bodily portion to which each pixel in depth pictures corresponds is identified by a label. Reliable offers for the locations of 3D cadaverous joints are the outcome of the algorithm for recognizing mortal disguises. A decision tree labels pixels in a rather noisy manner. A noisy labelled image built on mean shift offers the cypher common [23]. the use kinematic shell limitations to enhance the delicacy and robustness of typical offers. The severity of the bones and their fixed connection are used to infer the kinematic restrictions. A femur, for instance, connects a hip and knee joint and places restrictions on the range of motion for each joint. The depth values between the connected joints should fluctuate linearly from the knee to the hipsterism within an error threshold since they are intended to measure on the face of the ham. We can recognize body corridor mislabeling thanks to these restrictions [21]. The model was written in Python using the Keras Successional API. The shape of the input case is 45\*9\*18, which represents 45 success frames with 18 key points that each include X and Y coordinates. ReLU activation is having to key spots of every frame for point birth in time-distributed CNN subcaste with 16 contaminants of size 3\* 9\* 3. CNNs are better able to value scale- and gyration-steady spatial characteristics. The CNN subcaste can value spatial characteristics like the angles and distances between the vibrant focal points in a frame. The CNN affair is subjected to batch normalization for quicker convergence. This is a powerhouse subcaste that prevents overfitting by dropping some weights at random. The CNN data is applied to each of the 45 frames and smoothed before being fed with 20 units and a 0.5 unit forget bias to the LSTM sub-caste. The CNN subcaste uproots certain features, and the temporal changes in the features are recognized by LSTM [22]. A body posture modeling and evaluation system to celebrate and assess yoga poses and suggest instruction to learners is proposed in this study. A two-stage classifier was built using FCM and BP-ANN to model and celebrate full-body poses. Because of its crucial nonlinear processing, BP-ANN Detectors recommended as the first classifier to classify yoga stations into distinct orders capability, and FCM was espoused as the successive classifier to categorize the postures in a order for the flexible fuzzy partition. The two-stage classifier could handle the variations in subject position while improving recognition outcomes at a reasonable computational cost. Extension: To further improve the recognition delicacy, we also suggested a recognition technique to reject insulated honored products and noisy results with accretive probability [8]. The article suggests a system for correcting posture or poses. With

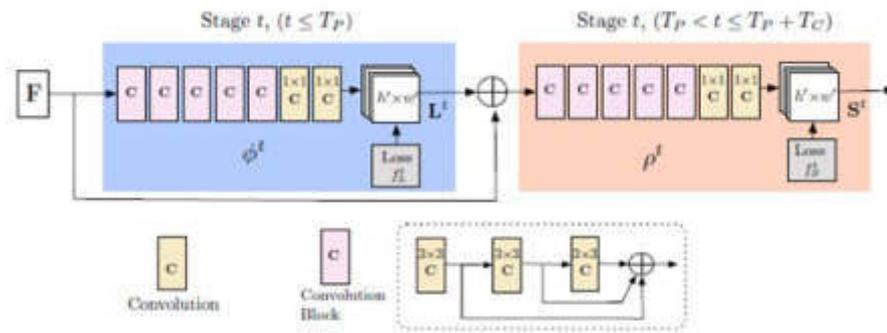


Fig. 2.1: Open Pose Architecture Model

YogaST, three categories of poses have been created using the kinect technology: Downward Dog, Tree Pose, and Warrior Pose are examples. A yoga expert system has been created using texts and graphics to direct the practitioner. the approximately 5500 photos in the YOGI collection [14]. Angles and joints from these photo frames were fed into various machine learning models, including Logistic Regression, Random Forest, and Support Vector Machine, using the "tf-posture" technique (SVM). In the experiment, skeletal characteristics of the human posture in the photo frame were extracted using Open Pose architecture, and they were then stored in a NumPy array. Its model is a straightforward neural network architecture with two hidden layers and an output layer [5]. The posture is estimated with a help of web camera that helps us in capturing the real time image. They took into consideration 12 sun salutation asana poses that were recorded on a webcam. The method creates a skeletonized image and on that image feature extraction is applied and through ML model a sun salutation is classified along with Support Vector Machine (SVM), k-nearest neighbors (KNN), Logistic regression and Naïve Bayes producing 96% of the result [9].

The VGG 19 convolutional neural network is the model utilized for the experiment. It comprises 19 deep layers that are made up of a successional composite of complication, corrected direct unit activation, and Max Pooling layers. illustrates the VGG-19 model's connected layers. The size of the input subcaste is 224 x 224 x 3, which corresponds to the height, range, and RGB (red, green, and blue) channel separately [2]. the interconnected successional layers that make up our final experimental model. Figure 8 shows the Transfer learning VGG 19 model's armature. A SoftMax activation function that conforms to five class markers for each of the a fore mentioned yoga acts makes up the affair subcaste. The Adam optimizer, which combines grade descent with incitement and root mean square propagation, is used in this model. Since it is a multiclass bracket model, the Adam optimizer's literacy rate is set to 0.001 and its loss function is categorical cross entropy [1]. The model's performance is covered with a batch size of 16 and several ages or passes set to 25. demonstrates the VGG19 model's performance standards on the train and test datasets. We notice a more suggestive change in the model's delicacy after five ages or passes, and it plateaus at 98 percent delicacy after twenty-five [13].

The suggested model's DCNN armature is like the models. The three layers (convolutional layers) along with the max-pooling layer with a 2x2 window size and of 2 strides make up the feature extraction portion. The first three convolutional layers comprise 32(first layer), 128(second layer), and 256(third layer) each with a size of 3x3 and 1 stride. Because there are three maximum pooling layers used in the armature, the point feature extraction creates 256-point maps that are three times smaller (3\*1) than the size of the input image (24\*8). likewise, the fully interconnected layers receive these qualities (the bracket part). The first, alternate, and third layers, which make up the bracket portion, respectively have 128, 64, and 32 neurons each and use the ReLU activation function. However, the final subcaste uses the SoftMax activation function and contains 26 neurons per labor. Since we used a tenfold cross-validation approach during the training procedure, ten DCNN models were trained for 100 ages [11]. This study employs Convolutional Neural Networks (CNNs) for the identification of various poses. Utilizing Kinect, the study investigates a range of image combinations, including RGB, Depth, RGB-D, and background-subtracted RGB-D [4].

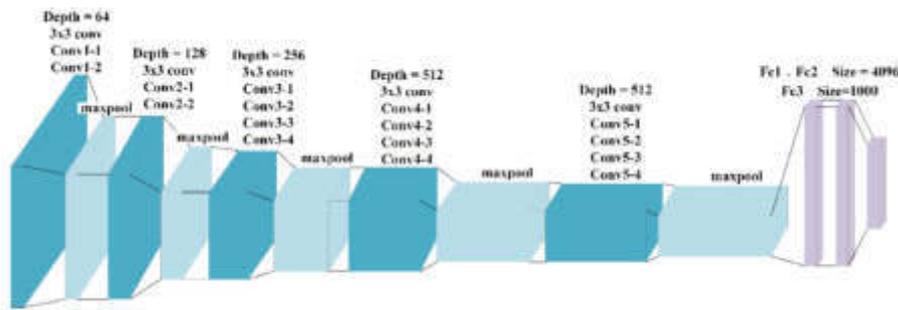


Fig. 2.2: VGG19 Architecture

A custom dataset was created by capturing diverse activities carried out by various individuals in various indoor environments. The findings indicate that the most effective approach for indoor video-based fall monitoring involves combining RGB background-subtracted and Depth with CNNs [3]. In this research paper, we introduce a system designed for recognizing Yoga postures. This system not only identifies the specific Yoga pose being performed by a practitioner but also retrieves relevant Yoga training information from the Internet to provide guidance on correct posture [15]. To achieve this, we initially employ a Kinect device to capture the user's body map and extract the body contour. Subsequently, we utilize 'star skeleton,' a rapid skeletonization technique that connects from the centroid of the target object to the contour extremes. This star skeleton serves as a representative descriptor for recognizing the practitioner's Yoga posture.

Lastly, the system retrieves Yoga training information associated with the recognized posture from the Internet, assisting the practitioner in maintaining the correct posture during their practice [20]. This paper introduces a system designed to track and assess the precision of various yoga poses, thereby assisting users in their yoga practice. The system utilizes Microsoft Kinect to capture real-time data on the movement of different body parts, allowing for the calculation of angles based on the detected joint points. These angles are then employed to evaluate the accuracy of specific yoga poses for the user. The system has demonstrated the capability to promptly identify and assess a variety of yoga poses in real-time [7]. This paper aims to construct a model for training data that aligns with human posture characteristics, thereby breaking down this complex problem to reduce computational complexity and enhance system performance in practical applications. Through real-world experimentation, the model identifies distinct body movement postures by observing sequences of human postures, employing matching, identification, and classification processes [19]. The results affirm the feasibility and effectiveness of the proposed method for human posture recognition. Furthermore, for detecting human movement targets, the paper introduces a method based on Gaussian mixture modeling, and for extracting human object contours, it proposes an approach based on the Sobel edge detection operator.

Additionally, this study includes an experiment on human posture recognition and an assessment of our cloud-based monitoring system for elderly individuals utilizing our methodology [18]. The proposed system begins by detecting humans before conducting posture analysis, restricting the posture recognition process to human silhouettes. The human detection technique is intentionally designed to exhibit resilience to diverse environmental factors. As a result, posture analysis relies on straightforward and efficient features that are intended for characterizing human silhouettes without considering environmental constraints [10]. The posture recognition method, employing fuzzy logic, can identify four static postures and is robust against variations in the camera-to-person distance and variations in a person's body shape. With a satisfactory posture recognition accuracy of 74.29%, this approach can effectively identify emergency situations, such as a fall, within a health-oriented smart home [17]. Visual Gesture Builder, a data-driven machine-learning tool for detecting gestures, was employed to accurately capture essential yoga movements. This gesture analysis technology is currently under investigation for potential integration into exergames designed for personalized medical interventions. The research objective is to evaluate whether a machine learning algorithm embedded in a basic computer-based video exergame can evaluate the acquisition of yoga skills in specific target populations, thereby promoting

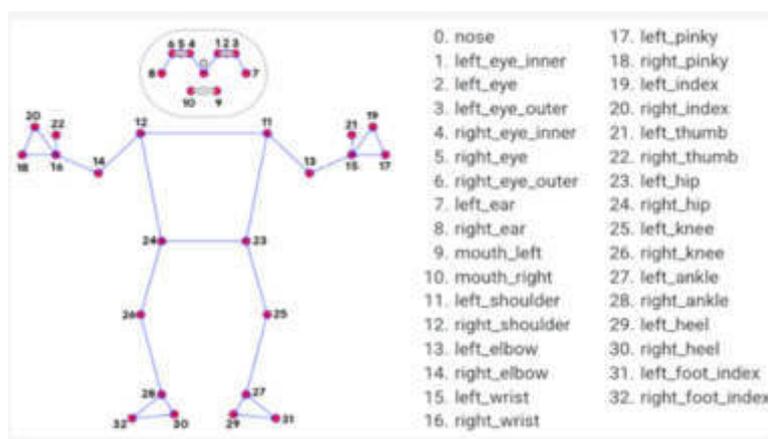


Fig. 3.1: Land Marks in Media Pipe Pose

healthy physical activity [6]. This paper introduces an alternative, computationally efficient method for recognizing Yoga poses in complex real-world settings using deep learning. In pursuit of this goal, a Yoga pose dataset was curated, featuring the participation of 27 individuals (comprising 8 males and 19 females) and encompassing ten Yoga poses: Malasana, Ananda Balasana, Janu Sirsasana, Anjaneyasana, Tadasana, Kumbhakasana, Hasta Uttanasana, Paschimottanasana, Uttanasana, and Dandasana. The videos were recorded using smartphone cameras with a 4K resolution and a 30-fps frame rate. For real-time Yoga pose recognition, a three-dimensional convolutional neural network (3D CNN) architecture was devised and implemented. This architecture is a modified version of the C3D architecture initially designed for human action recognition. In the proposed modified C3D architecture, computationally intensive fully connected layers were pruned, and additional layers like batch normalization and average pooling were introduced to enhance computational efficiency. To the best of our knowledge, this study is among the first to leverage the inherent spatial-temporal relationships among Yoga poses for recognition. The designed 3D CNN architecture achieved a test recognition accuracy of 91.15% on the in-house curated Yoga pose dataset containing ten distinct Yoga poses [16].

**3. Data Collection.** All the videos contain amassed for greater than 40 seconds in an enclosed space at a frame rate of 35 frames every second. An overall of 88 training videos are 1.5 hour, 5 minutes, and 5 seconds long at 35 frames every second, or roughly 750 frames. For every terminal state of each pose, roughly a hundred samples should be collected in order to create an effective classifier. It is critical that the samples used in the collection reflect a range of camera perspectives, landscape features, body types, and position variations. 33 pose landmarks are predicted by the MediaPipe Pose data set landmark model.

The data set body mark model in Media Pipe Pose predicts the area of overall pose landmarks nose, left\_eye\_inner, left\_eye, left\_eye\_outer, right\_eye\_inner, right\_eye, right\_eye\_outer, left\_ear, right\_ear, mouth\_left, mouth\_right, left\_shoulder, right\_shoulder, left\_elbow, right\_elbow, left\_wrist, right\_wrist, left\_pinky, right\_pinky, left\_index, left\_thumb, right\_thumb, left\_hip, right\_hip, left\_knee, right\_knee, left\_ankle, right\_ankle, left\_heel, right\_heel, left\_foot\_index, right\_foot\_index A full-body segmentation mask displayed as a two-class segmentation can be predicted by Media Pipe Pose (human or background).

## 4. Methodology.

**4.1. Neural Network Architecture.** This model majorly works with neural network where the model tries to imitate the human brain to recognize the yoga posture. Initially, after the skeletonization of the image or the real time image, the model grasps the random inputs from the input nodes and passes through the hidden layer providing the output. But before resulting in the output the model tries to recognize the detected model is relevant to its posture or not through the other proposed algorithms and goes through backpropagation method. As the human brain tries to learn from the errors, the model imitates the same format of human learning from

Table 3.1: Key points Used

No	Key Point	No	Key Point
0	Nose	17	left_pinky
1	left_eye_inner	18	right_pinky
2	left_eye	19	left_index
3	left_eye_outer	20	right_index
4	right_eye_inner	21	left_thumb
5	right_eye	22	right_thumb
6	right_eye_outer	23	left_hip
7	left_ear	24	right_hip
8	right_ear	25	left_knee
9	mouth_left	26	right_knee
10	mouth_right	27	left_ankle
11	left_shoulder	28	right_ankle
12	right_shoulder	29	left_heel
13	left_elbow	30	right_heel
14	right_elbow	31	left_foot_index
15	left_wrist	32	right_foot_index
16	right_wrist		

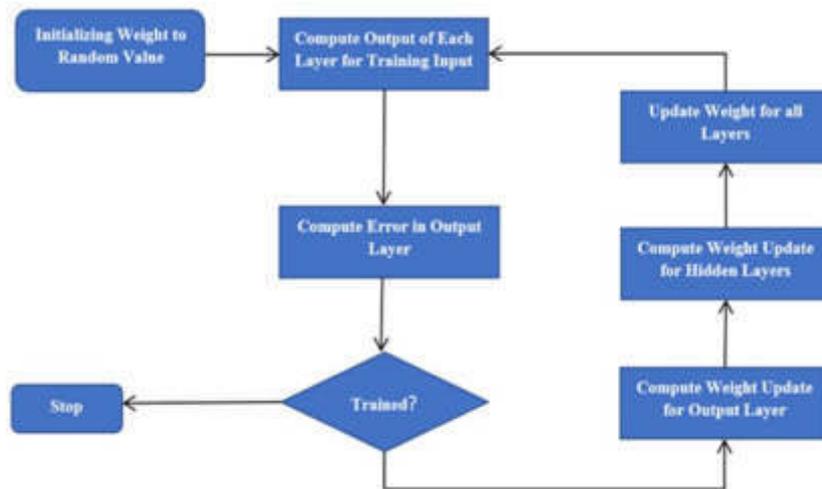


Fig. 4.1: Neural Network Architecture

errors and mistakes. The model goes through all these phases and recognizes the exact yoga posture and using media pipe the points allocated to each phase and part of the body, the model reads the skeletonized image and through the angles between the points allocated to joints and guides us in practicing the right posture.

A major portion of the exploration work on the Yoga Dataset has been carried out using Convolutional Neural Networks (CNN). A complication is to input as a sludge performing in activation. It can learn a huge number of pollutants under constraints specific to a training dataset. The armature substantially comprises colorful layers with pollutants and activation functions. The convolutional subcaste applies pollutants to the input image. Pollutants are vital in armature since they descry the spatial patterns in an image grounded on the change in intensity values of pixels. A pooling subcaste is added to reduce computational costs by

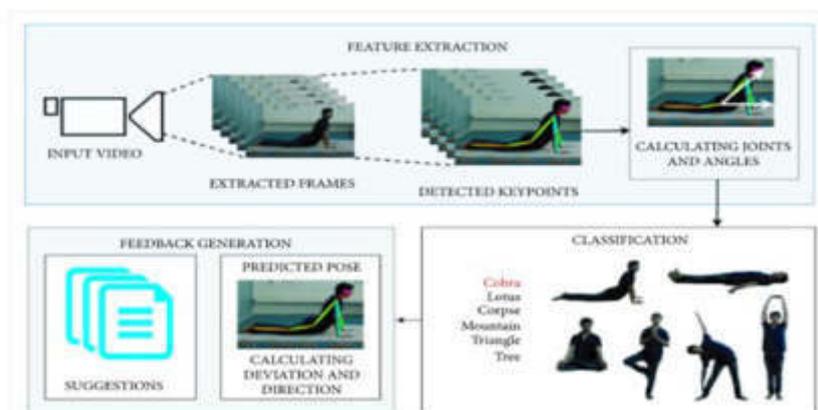


Fig. 4.2: Pose Extraction

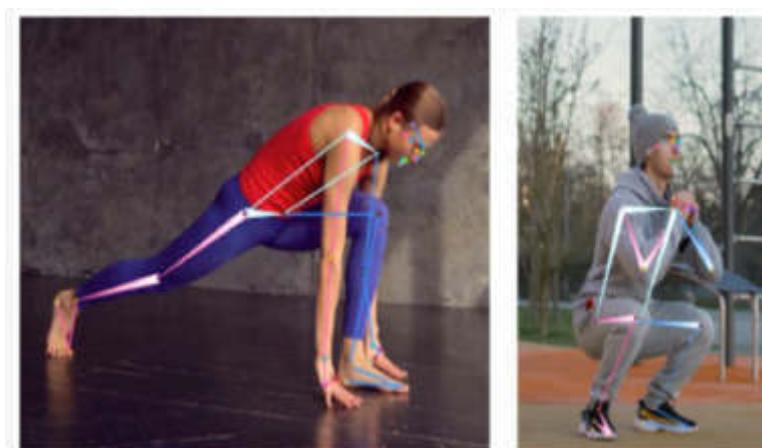


Fig. 4.3: Media Pipe Key points on Actual Humans

dwindling interlayer connections. Weights and impulses are added to completely connected layers, which connect neurons from different layers. Powerhouse layers are added to drop neurons, reducing the size of the model. Experimenters have experimented with different layers and activation functions in their work. Shows the armature of a Convolutional neural network.

From figure 4.2 describes about Keyframes are created from each frame of the videotape in the first phase, which is followed by their storage in JSON format. Key points are various body parts that make up a person and are important for the formation of a yoga disguise; examples include the shoulders, elbows, wrists, knees, etc. For important aspects, we used the Media Pipe library, a cross-platform toolkit created by Google that offers incredible ready-to-use ML results for computer vision applications.

A highly optimized, trained CNN model is utilized in this phase to infer thirty-three 3D milestones and entire body from RGB video tape masks with the background segmentation frames. This phase also uses a CNN model for high-dedication body position shadowing. Three equals are generated by the Media pipe library, with Z denoting the depth of a 2D match. (10). Fig. 3.1 displays the 33 key points provided by the Media Pipe library, whereas Fig. 4.1 shows the outcome of using the key points to root the data.

**4.2. Training.** In this paper, we have analyzed the yoga pose classification techniques in detail and have also created a model to contrast with the existing approaches. While most of the models exceed expectations

Table 5.1: Accuracy of Various Classes

S.No	Name	Class	Accuracy(%)
1	Validation	Framewise	98.45
2	Training	Framewise	97.56
3	Real Time	polled	99.05
4	Test	Polled	96.34
5	Test	Framewise	97.89

with higher performance metrics, the dataset needs to be more mature. An ideal dataset should be diverse and should reflect real life as much as possible. Overfitting and Underfitting a machine learning model can perform worse on newer data and leads to poorer accuracy and other performance metrics. Data Augmentation techniques can help in producing more data and thereby improve the size and the quality of the dataset For the visualization of asanas, features are carried out using Open Pose and the common position values are recorded in the JSON train. CNN and LSTM (Long Term Short Memory) models are also connected. Due to the combination of the two, we can recognize the fashionable set of data created by CNN (Convolution Neural Networks) and the data dependencies for long term created by LSTM. Theano's backend and Keras are used to collect the model. The performance of the entirely linked subcaste's relationship with SoftMax activation is scaled using the categorical because of its suitability cross-entropy loss function happened. Because of the straightforward inputs and small design, the training is quick at about 22 seconds every time. Over the course of training, both the modify in delicacy and the loss function independently. Initially, the accuracy of both increases fleetly with training accuracy staying below the validation accuracy The variations for training and confirmation accurateness, and training and validation misplacements throughout network training. The both accuracy increases steadily for the first 10 years, as depicted in the figure. however, a marked difference in accuracy can be seen in later ages. Also, it is possible to see that the validation accuracy detects the training accuracy, shows that sufficiently generic model is and will render the chosen Yoga poses coherently on test or unseen Yoga options. Once the network converges for the 25th time, none of the two measures—training accuracy nor validation accuracy—undergoes any substantial change (the model becomes pregnant). At the fifteenth repetition, the errors and misplacements transfer their asymptotic values. Nevertheless, When the model was impregnated, training was stopped after another 18 epochs, with the validation accuracy remaining unchanged. As a result, the Model Checkpoint call back job automatically saves the model of the weights corresponding to the time having good validation accuracy, for conclusion test data of yoga disguise input videos. In conclusion, our developed model successfully fitted the dataset of difficult real-world Yoga poses from our own set during training within 25 epochs.

**5. Results and Discussion.** The perpetration of other transfer literacy models for image classification similar as ResNet50, Inceptionv3 and Efficient Net. Advanced data addition ways can be used to give added quality- concentrated data which aids the model's interpretation. Although it is computationally precious, tentative GANs and neural style transfer styles transfigure an image from one sphere to another furnishing robust data to feed the neural network models. The count of acts used to develop in the data set to aggrandize the dataset's quality and bring further credibility to the use of the yoga classification model. Table 5.1 represents the accuracy of various classes.

The work provides self-train systems for recognizing and correcting the posture into consideration and uses the RGB webcam. In our data consisting the yoga postures based on the image of the system performed well and giving instructions to correcting the yoga pose includes totally various from each and every appearance, it gains an total accuracy of 90% in feature extraction.

They used homemade point birth and make a distant model for every yoga pose, which performs and gives lower time count devouring to apply and requires a good number of features add . In our route a good order of Yoga is reachable as can append any neuron in every t thick subcaste and again train the model on the substitute data in the dataset. The given model contains 99.38% is also prideful to containing model as shown

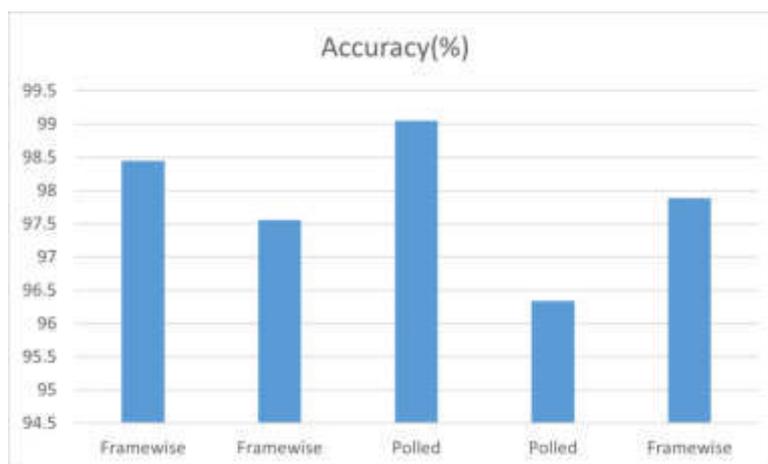


Fig. 5.1: Comparison of Accuracy for various classes



Fig. 5.2: Adho Mukha Svanasana

in figure 5.1.

**5.1. Real Time Yoga Posture.** From fig 5.2 Based on the trained data, the model suggests the person instructions to make the posture perfect. By comparing the real time practice to the trained practice with the body points and angles between the joints, the model has instructed the person to lift his right arm, extend the angle at right hip, and to extend the right arm at elbow giving the score 75.

From figure 5.3 this posture got some different angles compared to the previous yoga pose. These angles are now to be compared to the real time practice. As we can see the both arms of the person, the left arm is a bit tilted whereas the other arm is straight. The model instructed the person to lift his right arm and his left arm.

From figure 5.4 the person is instructed on lifting his left and right arm and extending and reducing the angle at his right and left hip to make the posture perfect giving the score 77

From figure 5.5 the current posture lacks some details from the actual posture and the person is instructed on extending his both right and left arms reducing the angles at his right hips and knees and extending the right leg more. According to the score 67, the person needs to work on this posture more.

From figure 5.6 the practitioner got advised about lifting his both arms to make the namaskar gesture



Fig. 5.3: Virabhadra Asana



Fig. 5.4: Adho Mukha Svanasana

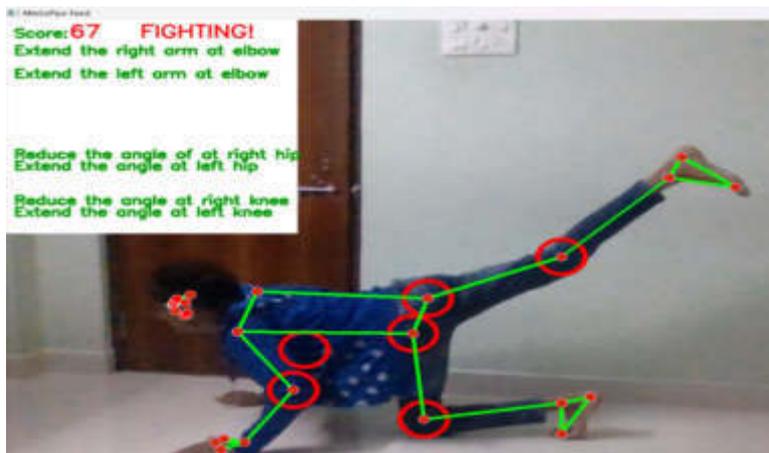


Fig. 5.5: Marjarya Bitil Asanaa



Fig. 5.6: Vrksasana Pose

straight and perfect and to make his right leg straight. And about the left knee, the person still needs to increase the space between the knee that is to increase the angle.

**6. Conclusion.** We have created a system in this design that consists of a channel for disguise identification, point localization on the mortal body, and an error identification method. This technique attempts to aid individuals in correctly yoga practice on their own and assist with ailments that may result from improper yoga poses. The approach using the algorithms CNN (Convolution Neural Networks) and the algorithm LSTM (Long Short-Term Memory) on disguise data attained to the library of Open Pose for Yoga actions discovery and set up to be largely successive. A device identifies the all poses on recorded videos as of the all activities of real time for every person. The devices gained detects Yoga shows in a videotape with 99.04% for framewise and 99.38% after polling process of each 45 frames. The network device scored 97.92% in real time activity of different people shows its capability to achieve easy more contents and conditions. The approach is suitable for evaluating the stoner's disguise from the front and providing feedback so that they can improve their yoga disguise using deep literacy techniques. The designed model is mounted atop a dashboard that was likewise made with stoners in mind It would be commented that our route predicates the need for any different technical tackle for Yoga pose adjustment and can be enforced on taken input from employing RGB camera. additional yoga postures, more datasets, and real-time image and video activity are all included. The device can be mounted on a mobile for tone-training and real-time forecasting. This work improves as proof of effort identification.

#### REFERENCES

- [1] K. ADHIKARI, H. BOUCHACHIA, AND H. NAIT-CHARIF, *Activity recognition for indoor fall detection using convolutional neural network*, in 2017 Fifteenth IAPR International Conference on Machine Vision Applications (MVA), IEEE, 2017, pp. 81–84.
- [2] U. BAHUKHANDI AND S. GUPTA, *Yoga pose detection and classification using machine learning techniques*, Int Res J Mod Eng Technol Sci, 3 (2021), pp. 13–15.
- [3] D. BRULIN, Y. BENEZETH, AND E. COURTIAL, *Posture recognition based on fuzzy logic for home monitoring of the elderly*, IEEE transactions on information technology in biomedicine, 16 (2012), pp. 974–982.
- [4] X. CAI, Y. GAO, M. LI, AND W. SONG, *Infrared human posture recognition method for monitoring in smart homes based on hidden markov model*, Sustainability, 8 (2016), p. 892.
- [5] A. CHAUDHARI, O. DALVI, O. RAMADE, AND D. AMBAWADE, *Yog-guru: Real-time yoga pose correction system using deep learning methods*, in 2021 International Conference on Communication information and Computing Technology (ICCICT), IEEE, 2021, pp. 1–6.
- [6] H.-T. CHEN, Y.-Z. HE, C.-C. HSU, C.-L. CHOU, S.-Y. LEE, AND B.-S. P. LIN, *Yoga posture recognition for self-training*, in MultiMedia Modeling: 20th Anniversary International Conference, MMM 2014, Dublin, Ireland, January 6-10, 2014, Proceedings, Part I 20, Springer, 2014, pp. 496–505.
- [7] B. DONAHOE-FILLMORE, C. J. BRAHLER, M. I. FISHER, AND K. BEASLEY, *The effect of yoga postures on balance, flexibility, and strength in healthy high school females*, Journal of Women's & Pelvic Health Physical Therapy, 34 (2010), pp. 10–17.

- [8] S. GARG, A. SAXENA, AND R. GUPTA, *Yoga pose classification: a cnn and mediapipe inspired deep learning approach for real-world application*, Journal of Ambient Intelligence and Humanized Computing, (2022), pp. 1–12.
- [9] M. GOCHOO, T.-H. TAN, S.-C. HUANG, T. BATJARGAL, J.-W. HSIEH, F. S. ALNAJJAR, AND Y.-F. CHEN, *Novel iot-based privacy-preserving yoga posture recognition system using low-resolution infrared sensors and deep learning*, IEEE Internet of Things Journal, 6 (2019), pp. 7192–7200.
- [10] R. R. GUDDETI, G. DANG, M. A. WILLIAMS, AND V. M. ALLA, *Role of yoga in cardiac disease and rehabilitation*, Journal of cardiopulmonary rehabilitation and prevention, 39 (2019), pp. 146–152.
- [11] M. U. ISLAM, H. MAHMUD, F. B. ASHRAF, I. HOSSAIN, AND M. K. HASAN, *Yoga posture recognition by detecting human joint points in real time using microsoft kinect*, in 2017 IEEE Region 10 humanitarian technology conference (R10-HTC), IEEE, 2017, pp. 668–673.
- [12] S. JAIN, A. RUSTAGI, S. SAURAV, R. SAINI, AND S. SINGH, *Three-dimensional cnn-inspired deep learning architecture for yoga pose recognition in the real-world environment*, Neural Computing and Applications, 33 (2021), pp. 6427–6441.
- [13] ———, *Three-dimensional cnn-inspired deep learning architecture for yoga pose recognition in the real-world environment*, Neural Computing and Applications, 33 (2021), pp. 6427–6441.
- [14] C. LONG, E. JO, AND Y. NAM, *Development of a yoga posture coaching system using an interactive display based on transfer learning*, The Journal of Supercomputing, (2022), pp. 1–16.
- [15] P. PULLEN AND W. SEFFENS, *Machine learning gesture analysis of yoga for exergame development*, IET Cyber-Physical Systems: Theory & Applications, 3 (2018), pp. 106–110.
- [16] G. J. SALEM, S. S.-Y. YU, M.-Y. WANG, S. SAMARAWICKRAME, R. HASHISH, S. P. AZEN, G. A. GREENDALE, ET AL., *Physical demand profiles of hatha yoga postures performed by older adults*, Evidence-Based Complementary and Alternative Medicine, 2013 (2013).
- [17] G. SATHYANARAYANAN, A. VENGADAVARADAN, AND B. BHARADWAJ, *Role of yoga and mindfulness in severe mental illnesses: A narrative review*, International journal of yoga, 12 (2019), pp. 3–10.
- [18] M. B. SCHURE, J. CHRISTOPHER, AND S. CHRISTOPHER, *Mind–body medicine and the art of self-care: teaching mindfulness to counseling students through yoga, meditation, and qigong*, Journal of Counseling & Development, 86 (2008), pp. 47–56.
- [19] J. K. SETHI, H. NAGENDRA, AND T. S. GANPAT, *Yoga improves attention and self-esteem in underprivileged girl student*, Journal of education and health promotion, 2 (2013).
- [20] T.-H. TAN, M. GOCHOO, S.-C. HUANG, Y.-H. LIU, S.-H. LIU, AND Y.-F. HUANG, *Multi-resident activity recognition in a smart home using rgb activity image and dcnn*, IEEE Sensors Journal, 18 (2018), pp. 9718–9727.
- [21] M. C. THAR, K. Z. N. WINN, AND N. FUNABIKI, *A proposal of yoga pose assessment method using pose detection for self-learning*, in 2019 International conference on advanced information technologies (ICAIT), IEEE, 2019, pp. 137–142.
- [22] Z. WU, J. ZHANG, K. CHEN, AND C. FU, *Yoga posture recognition and quantitative evaluation with wearable sensors based on two-stage classifier and prior bayesian network*, Sensors, 19 (2019), p. 5129.
- [23] K. YANG, K. YOUN, K. LEE, AND J. LEE, *Controllable data sampling in the space of human poses*, Computer Animation and Virtual Worlds, 26 (2015), pp. 457–467.

*Edited by:* Sathishkumar V E

*Special issue on:* Scalability and Sustainability in Distributed Sensor Networks

*Received:* Aug 31, 2023

*Accepted:* Nov 4, 2023



## INTERFACE CONTROL AND STATUS MONITORING OF ELECTRONIC INFORMATION EQUIPMENT BASED ON NONLINEAR DATA ENCRYPTION

MIN YAN\* AND HUA ZHANG†

**Abstract.** An advanced electronic information equipment interface control and status monitoring system is proposed to ensure the fairness, objectivity, and security of information while identifying responsibility for traffic accidents. Through an in-depth analysis of the system's security requirements and the current landscape of information security technology, a robust security strategy is developed for each crucial system stage. A PC-based platform is developed for efficient data acquisition, secure processing, reliable transmission, and fortified storage, focusing on implementing nonlinear data encryption methods. Performance evaluation of the system involved rigorous testing using files ranging from 3MB to 10MB. The results of the proposed system revealed a significant improvement in the system's overall speed and efficiency, showcasing an average performance enhancement of one quarter compared to the original platform. The proposed system demonstrated an impressive 15% to 30% increase in processing speed, establishing its capability to ensure data integrity protection during information transmission, facilitate accurate identification of data recording equipment post-accident, and safeguard the security of stored data. The developed electronic information equipment interface control and status monitoring system effectively addresses critical challenges associated with ensuring data integrity and security in traffic accident investigations.

**Key words:** Traffic accidents; Data encryption; Information security; Fairness; Objectivity

**1. Introduction.** In the current era of rapid economic growth, the exchange of information has become increasingly crucial. However, certain factors, such as the propagation of connecting cables and space saturation, significant challenges to efficient information transmission. To tackle these issues, scientists have developed various short-range wireless communication technologies. These technologies involve communication distances of less than 200 meters and employ radio technology for data transmission among different devices. Notable examples of such technologies include Bluetooth, Wireless Fidelity (WiFi), IrDA, UWB, Near Field Communication (NFC), ZigBee, WLAN, Ad Hoc, WMN, and ANT, each occupying distinct wireless frequency bands and offering unique performance, technical advantages, and application environments [123].

The significant advancements in Bluetooth's technical specifications, particularly with the release of the Bluetooth 4.0 specification in 2010, the competitiveness of Bluetooth products in the market has surged, leading to an expanding scope of applications. Statistical data indicates a rapid upsurge in the production of Bluetooth technology products since 2007, with global sales surpassing 1 billion units in 2008, reaching 1.6 billion units in 2011, and nearly 2 billion units in 2012. According to ABI's market tracking report on the Internet of Everything (IoE) in the second quarter of 2015 (MD-IOE-104), the worldwide count of Internet of Things interconnection devices is projected to exceed 45 billion by 2020, with at least 14 billion of these devices utilizing Bluetooth specifications. Bluetooth technology products are anticipated to account for at least one-third of Internet of Things interconnection devices by 2020. The widespread application and integration of Bluetooth products within the increasing Internet of Things landscape have made them ubiquitous in everyday life [456].

Despite being the primary wireless communication technology acknowledged by the industry, concerns regarding the security of Bluetooth technology have persistently echoed throughout the external sphere and the extensive user community. With the continuous evolution of network technology and the substantial enhancement of computer processing capabilities, Bluetooth's conventional encryption algorithm mode has become inadequate to meet the security demands of the contemporary application landscape. Presently, the utilization of Bluetooth applications has witnessed consistent spread. Notably, not only traditional sectors such as medical, commercial, financial, defense technology, aerospace, and the military, characterized by high information value and stringent

---

\*XinXiang Vocational and Technical College, School of Information Engineering, 453006, China ([minyan298@163.com](mailto:minyan298@163.com))

†Xinxiang Vocational and Technical College, Xinxiang, Henan, 453621, China ([huazhang58@126.com](mailto:huazhang58@126.com))

security prerequisites, impose fresh security requisites on Bluetooth, but also various mass industries, including the automotive sector, security, automated control, environmental monitoring, smart home technology, urban management, modern architecture, and the Internet of Things, have thrust Bluetooth security concerns to the forefront. Consequently, Bluetooth technology's current data security performance is causing significant apprehension [7, 8].

**2. Literature review.** In 1994, Ericsson's mobile communication department pioneered developing a cost-effective and energy-efficient wireless technology to facilitate the connectivity between mobile phones and nearby devices, eventually named Bluetooth. Subsequently, in February 1998, Ericsson, in collaboration with IT industry leaders including Intel, KM, Nokia, and Toshiba, established the "Bluetooth Special Interest Group" (Bluetooth SIG), formally solidifying its presence in the technological landscape. The consortium welcomed the participation of additional technology giants such as 3Com, Microsoft, Motorola, and Lucent, who joined the ranks of Bluetooth SIG as active proponents [9].

On July 26, 1999, Bluetooth SIG introduced Bluetooth specification 1.0, marking a significant milestone in the evolution of the technology. This was followed by the release of Bluetooth Specification 2.0 in November 2004, which proposed the Enhanced Data Rate (EDR) transmission mode. The latest Bluetooth specification, 3.0, was unveiled on April 21, 2009, incorporating the high-speed transmission mode, boasting a maximum 24 Mbps [10].

The concept of low-power Bluetooth, originally conceptualized by the Nokia Research Center, began its developmental journey in 2001 and was officially launched in October 2006 under the name Wi-Bree. Initially targeted at applications in personal medical and security domains due to its exceptional low-power Design and compatibility with traditional Bluetooth, Wi-Bree exhibited strong market potential in the medical and safety sectors. Consequently, the Wi-Bree Forum announced its incorporation into the Bluetooth Special Interest Group (SIG) on June 12, 2007, leading to its renaming as Ultra Low Power (ULP) Bluetooth [11].

In mid-2008, ULP was rebranded as Bluetooth Low Energy, with plans to introduce low-power Bluetooth specification products in mid-2009. Subsequent advancements led to the official launch of Bluetooth Specification 4.0 by the Bluetooth Special Interest Group on June 30, 2010, combining traditional, high-speed, and low-power Bluetooth capabilities, propelling Bluetooth technology to a new level. The Bluetooth Technology Alliance continued this trajectory with the introduction of the Bluetooth 5.0 specification on June 16, 2016, announcing substantial enhancements in transmission distance, speed, and data capacity, thereby setting the stage for a new generation of Bluetooth standards from the end of 2016 to the beginning of 2017 [12].

The author's primary focus centres on conducting an in-depth analysis of the system's security prerequisites and proposing comprehensive security measures for each system component, closely integrating insights from information security technology. Emphasizing the system design's pivotal aspects and pragmatic aspects, the author has developed a PC-centered platform dedicated to data gathering, secure processing, continuous transmission, and reliable storage. This platform validates the establishment of communication channels and the effectiveness of the data protection processes [13].

### 3. Research Methods.

**3.1. Safety Demand Analysis.** The primary risk to the secure transmission of data within the system during the data processing phase arises from potential transmission failures or wireless interferences during the data transfer between the front-end acquisition unit and the central unit. These challenges concerning data integrity, timely delivery, and the security of Bluetooth wireless connectivity. Furthermore, ensuring storage security involves verifying the freshness and integrity of data upon its arrival at the central unit. The goal is to guarantee the authenticity of device identity and data integrity, thus necessitating implementing the proposed storage process security measures. The safety structure diagram showing the data processing process within the system is presented in Figure 3.1 [14].

The fundamental security of the system is chiefly responsible for safeguarding the integrity and protection of the system's software and hardware elements, alongside the associated technologies utilized in their creation and operation. This entails implementing robust strategies to reinforce the inherent defenses of the system against potential vulnerabilities and threats. Notably, it emphasizes the deployment of dependable protocols and mechanisms to strengthen the system's flexibility against unauthorized access and breaches.

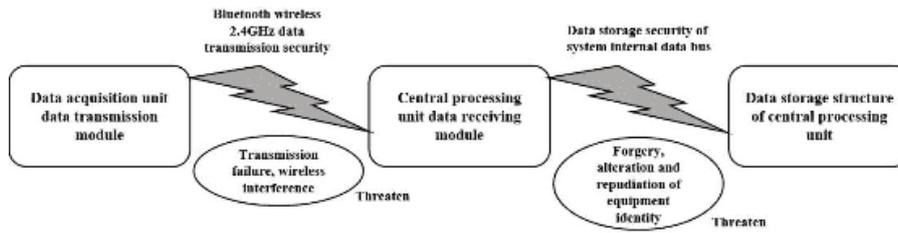


Fig. 3.1: Safety structure of system data process

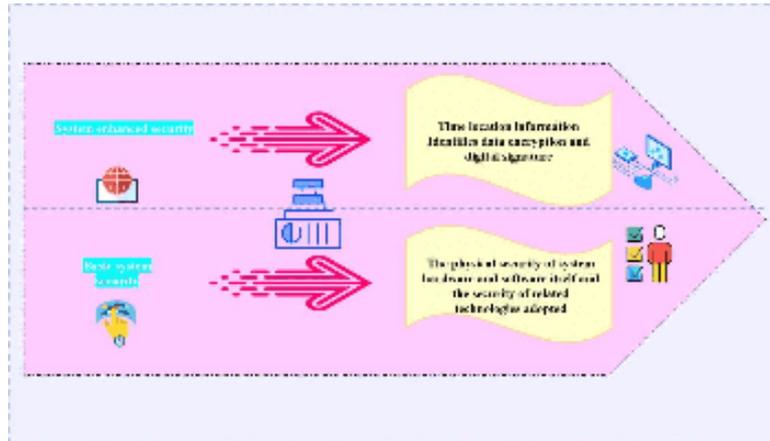


Fig. 3.2: Schematic of system security

Conversely, the heightened security dimension of the system primarily revolves around the meticulous management of data, incorporating sophisticated information security protocols and cryptographic methodologies to establish an additional layer of defense. The system uses sophisticated cryptographic techniques to encrypt and decrypt sensitive data, reinforcing its capacity to resist potential cyber threats and unauthorized breaches.

The system’s security framework, encompassing both the foundational and advanced security layers, is concisely represented in Figure 3.2. This visual representation underscores the comprehensive structure of the system’s security architecture, highlighting the pivotal interplay between foundational security measures and state-of-the-art technical procedures for ensuring robust data protection and overall system integrity.

### 3.2. System security design and implementation.

**3.2.1. Data integrity design and implementation.** Within this system, the preservation of data integrity pertains to the seamless transmission of data from the front-end unit to the back-end unit via Bluetooth communication. Additionally, it ensures the perpetual authenticity of the recorded information before and after the back-end unit ultimately stores the data. This entails guaranteeing that the data remains unaltered throughout these two critical data transmission and storage stages [15].

In the area of cryptography, the concept of the hash function is intricately intertwined with safeguarding data integrity. Typically employed to generate concise data "fingerprints," the hash function aids in distinguishing data through its distinctive characteristics. Notably, any alterations to the data trigger corresponding changes in its associated "fingerprint." Consequently, in cases where data integrity is potentially compromised during transmission or storage, data integrity verification can be facilitated by recomputing the data’s fingerprint and cross-verifying it against the original data fingerprint. This process effectively ensures the continual authenticity of the data [16].

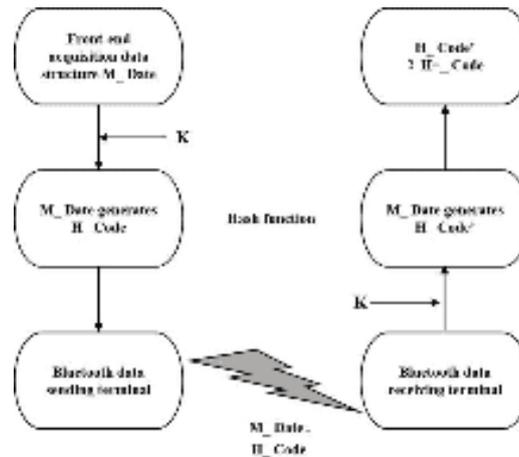


Fig. 3.3: Process for Ensuring Data Integrity

Let's consider the data structure *M\_Image*, which encompasses both temporal and spatial information stored within the system. The hash function, denoted as *H*, is employed in the context of the security mechanism implemented within the front-end unit. Notably, the processor within the front-end unit undertakes the hashing operation, as depicted in the following Equation 3.1.

$$H\_Code = H(M\_Image) \quad (3.1)$$

Consequently, the resultant output is the fingerprint of the recorded data, often referred to as the message digest of the data. Subsequently, during the transmission of the data structure package within the system via Bluetooth or upon the data's entry into memory through the system bus, it becomes necessary to re-execute the hashing operation exclusively at the receiving end of the transmission channel or when retrieving the data structure from memory. This operation is demonstrated as follows:

$$H\_Code' = H(M\_Image) \quad (3.2)$$

This results in the new condensed message representation of the data. Comparing *H\_Code* and *H\_Code'* for equality signifies the data's integrity. It is crucial to note that this determination relies significantly on the security attributes of the hash function. Moreover, given the embedded nature of the system components and the inherent constraints on data transmission and processing capacity, adopting a relatively uncomplicated and convenient hash function is imperative.

Specifically, the hash functions can be categorized into those without keys and those with keys. In cases where hash functions without keys are utilized, it is essential to ensure the secure storage or transmission of the message digest (*H\_Code*). Consequently, a hash function with a key to generate the condensed message summary is recommended. The outlined process is as follows:

- (1)  $H\_Code = H(M\_Image)$  and  $H\_Code' = H(M\_Image)$
- (2)  $H\_Code \neq H\_Code'$

Here, *K* represents the key of the HMAC function. In the practical data communication framework of the system, *K* functions as a provisional key solely dedicated to upholding the integrity of data transmission. Its primary objective is to secure the successful exchange of validation data between the front-end and back-end through the HMAC function. Consequently, the session key can be designated for both communicating parties during data transmission. It is crucial to securely store *K* within the system, ensuring its confidentiality and safeguarding its integrity. Figure 3.3 illustrates the design process for data integrity protection [17].

Within the data integrity mechanism framework, if the newly computed message digest differs from the initial message digest dispatched by the front-end unit, the system will discard the data. To ensure data integrity, the system utilizes a retransmission strategy to compensate for any discarded data.

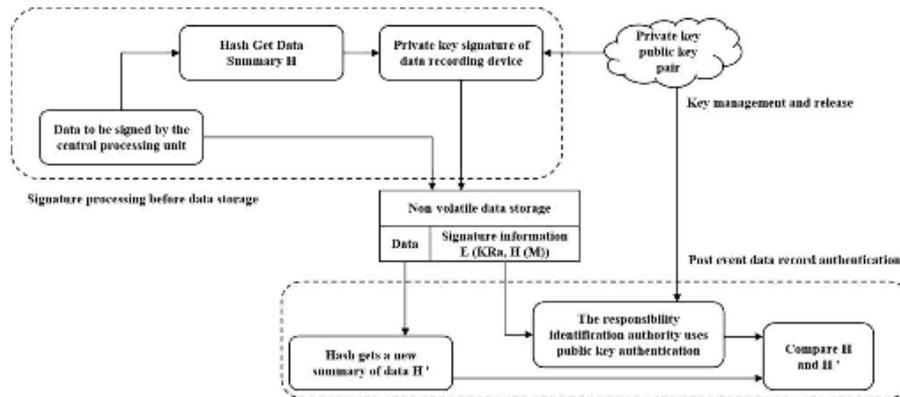


Fig. 3.4: Model for Implementing Digital Signatures using Public Key Cryptosystems

**3.2.2. Design and implementation of non-repudiation of data records.** Non-repudiation prevents the originator of a message from disowning the disseminated or transmitted message. In the system context, the vehicle monitoring data pertains to the information captured by the onboard equipment installed within the vehicle. As this data constitutes authentic records of the accident scene, it may contain information not in favour of certain parties. Consequently, to uphold the principle of impartiality, when the "evidence" is subsequently presented to the accident responsibility adjudication authority, the authority must verify the identity of the data recorder, i.e., the vehicle equipped with the monitoring system, and ascertain the origin of the evidence data. This verification is essential to prevent any evasion of unfavourable records. The non-repudiation mechanism within the system ensures the accountability of the data recorder and the data, thereby safeguarding the integrity and reliability of the entire process.

More specifically, the system's data security protection entails ensuring the security of the data both before and after storage. This comprehensive security measure is augmented by implementing a sophisticated digital signature scheme, representing an advanced level of protection within the system.

The data information the system captures incorporates the timestamp associated with the data recording process. In this context, verifying the consistency between the data generation time and the time of the recorded events is instrumental in ensuring the data's current state. However, this verification guarantees the data's chronological accuracy and does not provide comprehensive assurance of its authenticity. The system must employ a digital signature mechanism to guarantee non-repudiation of the data's origin [18].

In the system, attaining non-repudiation through digital signature for data relies on applying public key cryptography as an encryption algorithm. This technique employs a pair of keys, where one key is used for encryption, and another distinct yet related key is employed for decryption. Its notable advantage lies in the computational infeasibility of determining the decryption key solely based on knowledge of the encryption algorithm and key. When the public key encryption is utilized for digital signatures, the private key is initially employed for encryption, a process known as signing. To enhance efficiency, it may process either a segment of the message or the message summary.

As previously discussed, following the wireless transmission of data via Bluetooth and its verification for integrity, the data is stored in the temporary memory of the central control and data centralized processing unit. Before direct storage, the original data, which is currently devoid of any association with the recorder's identity, requires further security processing. The message digest of the original data, generated during the integrity verification, serves as the security hash code output by the corresponding data hash function. In this scenario, the system's private key encrypts this hash code to create a signature. Consequently, the recorded data concerning vehicle accidents and the accompanying signed information are stored in the non-volatile memory of the central control and data centralized processing unit. Concurrently, the system's signature represents the public key generated with the private key and disseminated through alternative channels such as the Public Key Infrastructure (PKI) system [19].

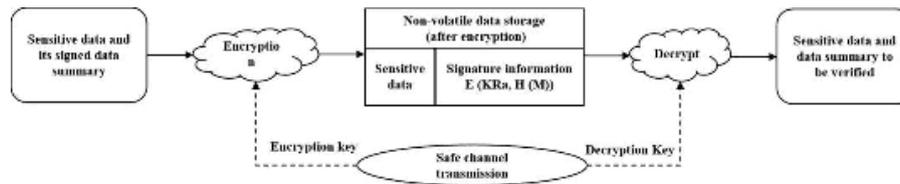


Fig. 3.5: Schematic diagram of data confidentiality processing

When the system's data necessitates submission to the certification authority for verification, the certifying entity extracts "evidence" data from the system's non-volatile storage. It validates it using the data record claimant's public key. This process verifies the authenticity and validity of the data recorder's identity, thereby facilitating the objective of responsibility identification. Figure 3.4 illustrates the schematic diagram of the implementation model, demonstrating the application of public key cryptography as a digital signature within the system. To verify the identity of the digital signature to the data recording device, here we use the RSA algorithm based on the mathematical problem of large prime number decomposition.

**3.2.3. Data confidentiality requirements.** Based on the preceding examination of the system's security requisites, it is evident that the comprehensive security framework for monitoring vehicle operation and accident data primarily emphasizes the integrity of wireless data transmission and the authentication of the data storage device. Typically, the data stored within the system is either inaccessible under normal circumstances or possesses minimal confidentiality demands. However, considering potential functional expansions, sensitive data such as GPS coordinates can be subjected to encryption.

In practice, the information logged by the onboard system doesn't require confidentiality measures for external parties. Even if the data is intercepted, the integrity of the key data recording information and the originating equipment remains unchanged. Consequently, the data retains the authentic details of the involved parties and upholds its legal validity as "evidence" for the responsibility identification authority.

In light of this, we establish the system's data confidentiality requirement as an extended security mechanism. The fundamental implementation concept involves the nonlinear data encryption after the signing process before its storage. Conversely, before its use as "evidence," the data undergoes decryption to eliminate the confidentiality shield, followed by public key verification. This approach is depicted in Figure 3.5.

In the data security design, we employed the AES symmetric encryption system. While ensuring the system's security demands, we have considered the performance criteria on the embedded hardware platform. Encrypting specific, limited-sensitive data fulfils the security needs and alleviates the functional strain on the system.

#### 4. Result analysis and Discussion.

**4.1. Test platform for Bluetooth transmission and data security processing.** Two key technological facets underpin the wireless Bluetooth vehicle safety monitoring system conceptualized by the project. Firstly, Bluetooth wireless data transmission's implementation and security technology facilitates data integrity protection, signature verification of data publishing devices, and encrypted storage of data message digests. To demonstrate the efficacy and practicality of these technologies within the system, the author has established a test environment on a PC platform, emphasizing Bluetooth wireless transmission and data security processing. This initiative has preliminarily actualized the research objectives and key technologies outlined in the system design.

To underscore the realization of Bluetooth communication and data security processing, the embedded control PC in the test platform has been substituted with a PC. The Bluetooth module is connected to COM1 and COM4, respectively. The software environment for the test platform has been compiled using Visual C++v6.0 and is executed accordingly. The two Bluetooth devices within the test platform adhere to the agreed-upon protocols for seamless communication.

Main equipment address: 00 80 38 14 3c CC COM1

Slave address: 00 80 38 14 3d 0f COM4

**4.1.1. Test platform software.** The testing software employs the Visual C++v6.0 serial port control, MSComm, to regulate the serial data communication interface. It effectively manages the Bluetooth module and facilitates data transmission via the HCI interface of the Bluetooth host.

**4.1.2. Bluetooth communication function realization.** Initialize the slave device (00 80 37 14 3d 0f): Select the function reset by chip, set the authentication enable, set the event filter, read the device address and query enable.

```
#define HCI_RESET                "01 03 0C 00"
#define HCI_WRITE_AUTHENTICATION_ENABLE    "01 20 0C 01 00"
#define HCI_SET_EVENT_FILTER        "01 05 0C 03 02 00 02"
#define HCI_READ_BDADDR            "01 09 10 00"
#define HCI_WRITE_SCAN_ENABLE      "01 1A 0C 01 03"
```

Bluetooth module reply time grouping.

At this time, the slave device has entered the waiting query state.

Main device initialization (00 80 37 14 3c cc): select the function by chip reset, set authentication enable, set event filtering, read device address and search device in sequence.

```
#define HCI_RESET                "01 03 0C 00"
#define HCI_WRITE_AUTHENTICATION_ENABLE    "01 20 0C 01 00"
#define HCI_SET_EVENT_FILTER        "01 05 0C 03 02 00 02"
#define HCI_READ_BDADDR            "01 09 10 00"
#define HCI_INQUIRY                "01 01 04 05 33 8B 9E 06 00"
```

Subsequently, the primary device yields the search outcomes, identifying a nearby Bluetooth device as the target slave device. Upon selecting the "create ACL link" function by the primary device, both the master and slave devices confirm the successful establishment of the link, denoted by the event group comprising the other device's address. The command structure for transmission is as follows: m\_StrAddress represents the location of the Bluetooth device's address, specifically the slave device address.

```
HCI_CREATE_CONNECTION1+m_strAddress+HCI_CREATE_CONNECTION2
```

Following the successful establishment of the ACL, the master-slave device events are synchronized accordingly.

Post ACL link establishment, the transmission of data files becomes feasible. The sender showcases the time and location details of the file's record and transmits it to the receiving end. Concurrently, the HMAC function is utilized to compute the summary of the data file. The function for calculating the data file summary is as follows:

```
void CBluetoothDlg:Mac(CByteArray & data, CByteArray & mac)
```

Assuming the HMAC function key is: "01234567890123456789," the data file information slated for transmission aligns with the information obtained by the receiving end.

Based on this, we can deduce that the sequence of message digests computed by the identical hash function with the key remains consistent before and after the file transmission. This consistency serves as a verification of the file's integrity throughout the transmission process.

**4.1.3. Implementation of digital signature and verification.** This section primarily includes two key functions: data security processing and security verification, enabling private key signature, data encryption and decryption, and public key verification. Depending on specific test requirements, signature and encryption can be independently chosen.

The public key information is manually inputted for device verification during the test without a Public Key Infrastructure (PKI) system for providing public key release. Furthermore, to ensure the security of the signature private key, the USBKey facilitates the private key signing in the RSA public key encryption algorithm.

The electronic key eKey is the "electronic key and ID card" within the client (i.e., the signatory). This USB-based product features a smart chip and a read-write controller, exhibiting a sleek design for convenient portability across various computers equipped with USB interfaces. It functions as a tool for digital signatures, identity authentication, secure information and data encryption, and the storage of personal certificates within a network environment. With a user storage space of 64K, it enables personal information, passwords, keys, certificates, and more storage. Additionally, it supports the RSA 1024-bit public key cryptography algorithm, allowing direct generation of RSA key pairs within the chip and facilitating functions such as signature/authentication and encryption/decryption. It provides an API dynamic library for streamlined programming.

Table 4.1: Comparison Results of Speed Test

File type	File size	Processing time(s)	
		Original platform	Author's text platform
MP3	3.10MB	7	7
EXE	6.35MB	13	9
ZIP	10MB	400	340

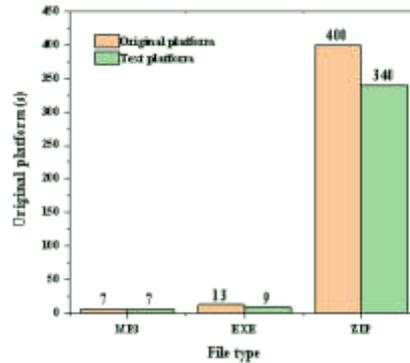


Fig. 4.1: Comparison of Processing Times between Original and Author's Platforms

During security processing, the initial step involves selecting the desired file for processing. If encryption is necessary, the encryption key must be entered. After selecting "Execute," the program prompts the user to store the file following security processing for verification purposes.

**4.2. Speed test.** Files within the range of 3MB to 10MB were chosen for testing purposes, and the outcomes of the tests are presented in Table 4.1 and Figure 4.1.

The Table 4.1 presents the processing durations for various file types on both the original and author's platforms. It offers insights into the processing time in seconds for MP3, EXE, and ZIP files, along with their corresponding sizes. The data emphasizes the superior efficiency of the author's platform, showcasing reduced processing times compared to the original platform for all file types. As indicated in Table 4.1, the author's platform demonstrates an average speed improvement of one-quarter over the original platform, resulting in an increased speed of 15% to 30%.

**5. Conclusion.** The integration of wireless Bluetooth technology with advanced data security processing, including data signature, encryption, and decryption, has led to the development model for a vehicle-mounted monitoring system. With the support of embedded control technology, this model maintains data integrity during information transmission, safeguarding against breaches and tampering for dependable data transfer. Furthermore, the system's ability to identify data recording equipment post-incident supports its security infrastructure, ensuring a robust mechanism for secure data storage. This contributes significantly to maintain the integrity of critical information providing credible evidence for post-incident analysis and investigations. Through the whole integration of these technologies, the system supports data transmission against potential vulnerabilities and establishes a strong foundation for ensuring data credibility and security in vehicle safety monitoring. As a robust tool, it guarantees sensitive information's accuracy, authenticity, and confidentiality, significantly enhancing safety and security standards within vehicle monitoring and surveillance.

## REFERENCES

- [1] Al-Absi, M. A., Al-Absi, A. A., & Lee, H. J. (2020, February). Comparison between DSRC and other short range wireless communication technologies. In 2020 22nd International Conference on Advanced Communication Technology (ICACT) (pp. 1-5). IEEE.

- [2] Zekavat, P. R., Moon, S., & Bernold, L. E. (2014). Performance of short and long range wireless communication technologies in construction. *Automation in construction*, 47, 50-61.
- [3] Vidakis, K., Mavrogiorgou, A., Kiourtis, A., & Kyriazis, D. (2020, June). A comparative study of short-range wireless communication technologies for health information exchange. In *2020 International conference on electrical, communication, and computer engineering (ICECCE)* (pp. 1-6). IEEE.
- [4] Joh, H., Yang, I., & Ryoo, I. (2016). The internet of everything based on energy efficient P2P transmission technology with Bluetooth low energy. *Peer-to-Peer Networking and Applications*, 9, 520-528.
- [5] Nusrat, T., Dawod, F. S., Islam, T., Kunkolienker, P., Roy, S., Rahman, M. M., ... & Braaten, B. D. (2022). A comprehensive study on next-generation electromagnetics devices and techniques for internet of everything (IoE). *Electronics*, 11(20), 3341.
- [6] Tyagi, A. K., & Nair, M. M. (2020). Internet of Everything (IoE) and Internet of Things (IoTs): Threat Analyses, Possible Opportunities for Future. *Journal of Information Assurance & Security*, 15(5).
- [7] Rege, K., Goenka, N., Bhutada, P., & Mane, S. (2013). Bluetooth communication using hybrid encryption algorithm based on AES and RSA. *International Journal of Computer Applications*, 71(22).
- [8] Albahar, M. A., Olawumi, O., Haataja, K., & Toivanen, P. (2018). Novel hybrid encryption algorithm based on aes, RSA, and twofish for bluetooth encryption.
- [9] Decuir, J. (2014). Introducing bluetooth smart: Part ii: Applications and updates. *IEEE Consumer Electronics Magazine*, 3(2), 25-29.
- [10] Zanella, A. (2009). A mathematical framework for the performance analysis of Bluetooth with enhanced data rate. *IEEE Transactions on Communications*, 57(8), 2463-2473.
- [11] Abdelatty, O., Chen, X., Alghaihab, A., Wentzloff, D. (2021). Bluetooth Communication Leveraging Ultra-Low Power Radio Design. *Journal of Sensor and Actuator Networks*, 10(2), 31.
- [12] Lin, J. R., Talty, T., & Tonguz, O. K. (2015). On the potential of bluetooth low energy technology for vehicular applications. *IEEE Communications Magazine*, 53(1), 267-275.
- [13] Tikkinen-Piri, C., Rohunen, A., & Markkula, J. (2018). EU General Data Protection Regulation: Changes and implications for personal data collecting companies. *Computer Law & Security Review*, 34(1), 134-153.
- [14] Huang, L., Wu, C., Wang, B., & Ouyang, Q. (2018). Big-data-driven safety decision-making: a conceptual framework and its influencing factors. *Safety science*, 109, 46-56.
- [15] Wang, H., Zu, B., Zhu, W., Li, Y., & Wu, J. (2022). On the Design and Implementation of the External Data Integrity Tracking and Verification System for Stream Computing System in IoT. *Sensors*, 22(17), 6496.
- [16] Hsiao, H. I., & Lee, J. (2015). Fingerprint image cryptography based on multiple chaotic systems. *Signal Processing*, 113, 169-181.
- [17] Garg, N., Bawa, S., & Kumar, N. (2020). An efficient data integrity auditing protocol for cloud computing. *Future Generation Computer Systems*, 109, 306-316.
- [18] Inam, S., Kanwal, S., Zahid, A., & Abid, M. (2020). A novel public key cryptosystem and digital signatures. *European Journal of Engineering Science and Technology*, 3(1), 22-30.
- [19] Khan, S., Luo, F., Zhang, Z., Rahim, M. A., Ahmad, M., & Wu, K. (2022). Survey on issues and recent advances in vehicular public-key infrastructure (VPKI). *IEEE Communications Surveys & Tutorials*, 24(3), 1574-1601.

*Edited by:* Venkatesan C

*Special issue on:* Next Generation Pervasive Reconfigurable Computing for High Performance Real Time Applications

*Received:* Aug 17, 2023

*Accepted:* Nov 2, 2023



## DETECTION AND PREVENTION OF CYBER DEFENSE ATTACKS USING MACHINE LEARNING ALGORITHMS

YONGQIANG SHANG\*

**Abstract.** Recent advancements in computing power, memory capacities, and connectivity have led to a corresponding surge in the utilization of big data, online platforms' prevalence, and machine learning's sophistication. Concerns regarding safety and the need for state-of-the-art security tools and methods to counter evolving cybercrime are amplified by the swift digitization of the world. This study investigates defensive and offensive applications of machine learning in cybersecurity. Additionally, it explores potential strategies to mitigate cyberattacks on machine learning models. The focus is on how machine learning facilitates cyberattacks, including developing intelligent botnets, advanced phishing using spear techniques, and deploying stealthy malware. Furthermore, the paper highlights the significance of artificial intelligence in digital safety, emphasizing its role in malware analysis, network vulnerability assessment, and threat prediction.

**Key words:** Internet risk estimation, Machine learning, Threat detection, Cyber safety, App-level protection

**1. Introduction.** Cryptography remains a top priority as the world transitions to a digital infrastructure. With advancements in network infrastructure, such as the World Wide Web, gaining access to cutting-edge science and technological breakthroughs has never been more convenient. Freeware and academic articles are increasingly accessible to the general public online. However, security experts and malicious players have easy access to state-of-the-art technology and research, each with their motivations for manipulating them. Safety measures have benefited from the development of techniques and innovations made possible by studies and advances in artificial intelligence. These technologies simplify the identification and effective counteraction of potential security threats. But, hackers can leverage this data to plan and execute more sophisticated and widespread attacks. Hackers hold a significant advantage in the ongoing battle on the internet, as they must only be successful once out of numerous attempts [1].

Achieving a 100% success rate is imperative when the goal is safety. Numerous studies revealed that in 2017, hackers targeted a diverse array of businesses, individuals, and mobile applications. Documents such as surveillance records, accounting information, and personal data were among the stolen information. The potential consequences of this data falling into the wrong hands, whether the general population or the black market, could be catastrophic. Here are some research findings on the impact of cybersecurity on businesses, organizations, and individuals [2]:

- Recently, over 350 trillion dollars have been lost or stolen due to cybercrime, including the costs incurred for repairing damages caused by criminal activities.
- A shortfall of over 1.8 million cybersecurity professionals is projected by 2022.
- Companies must spend at least \$100 billion annually to keep up with evolving technology.
- Intruders generate over a billion dollars yearly through ransomware attacks like WannaCry and CryptoWall.

Keeping step with and combating the escalating complexity of intrusions is becoming increasingly challenging, given the rapid obsolescence of safeguards. The average time to identify an intrusion is approaching seventy days. The growing magnitude and intricacy of intrusions further complicate efforts to keep up with the continuously emerging new threats and vulnerabilities. This challenge arises from the increasing sophistication and breadth of intrusions. Here, the authors focus on artificial intelligence, an emerging discipline with far-reaching implications for online safety [3].

---

\*Xinyang Agriculture and Forestry University, Department of Information Engineering, Xinyang Henan 464000, China ([yongqiangshang8@163.com](mailto:yongqiangshang8@163.com))

**1.1. Defence against cyber-attacks using machine learning .** Machine learning is a subset of AI that enables machines to learn and improve from data without being programmed. It employs mathematical frameworks derived from the analysis of patterns in datasets to achieve this objective. These models are then applied to make predictions based on newly acquired information. For example, machine learning systems are utilized in the e-commerce sector to customize product recommendations for individual shoppers. Moreover, machine learning is finding applications in the medical field for predicting outbreaks and assessing whether a patient is predisposed to developing cancer based on their medical history. Machine learning holds considerable potential for diverse applications across various fields [4].

Machine learning techniques are broadly categorized into prospective (controlled) artificial intelligence methods and pattern identification (or uncontrolled) machine learning techniques. In controlled learning, an algorithm is provided with a set of training data and tasked with predicting the value of a target variable using various learning methods. For example, a machine learning model may determine whether an internet protocol (IP) address is involved in a distributed denial of service (DDOS) assault based on factors such as the geographical origin of the IP, the frequency of web requests, and the time of day the requests were made. Various machine learning methods, including linear and logistic regression, decision trees, and support vector machines (SVMs), fall under the umbrella of supervised learning. Conversely, the objective of unsupervised learning is not to predict a specific variable. Uncontrolled algorithms learn to uncover relevant patterns and correlations in datasets. For instance, clustering and association algorithms may be employed to identify groups of malicious software with common operations and behaviour [5].

The domain of cybercrime and security is rapidly adopting machine learning. Beyond being a prevalent application area, the IT sector offers various potential uses, including malware analysis and log analysis, where machine learning is widely deployed. Both well-intentioned users and those with malicious intent leverage the powerful machine-learning capabilities of the internet. The upcoming section will explore the dual nature of machine learning's application in internet security and criminal activities [6].

**1.2. Machine Learning Techniques for Detecting Cyber Attacks.** Two approaches to cyber threat detection are signature and anomaly-based, employing ML techniques. Recent advancements in artificial intelligence methods have facilitated the establishment of signs to identify the code and behaviour of spyware accurately. These signatures are integrated into the construction process. Several strategies have been developed for the swift and efficient retrieval of signatures from versatile parasites, including NSG and LSEG, which are methods for creating signatures based on network behaviour. The LESG method primarily focuses on viruses propagated through buffer overflow attacks. Figure 1 illustrates the attack vectors in the world [7].

The F-Sign method extracts a signature from the worm's code, which can be utilized to detect and halt the worm's propagation. Alternative methods for generating signatures of malicious software based on the network traffic it produces include semantic aware (SA), as documented in scholarly publications. These technologies can accurately detect malicious activity even when the system has considerable noise. Building a model using anomaly-based approaches is a common practice for identifying cyberattacks, encompassing typical and typical network activity patterns [8]. These methods commonly employ unsupervised, semi-supervised, and supervised algorithms, all rooted in artificial intelligence. Techniques like k-means, fuzzy c-means, QT and SVM are frequently utilized for constructing clustering approaches in unsupervised learning. When these techniques aggregate network activity, a decision often arises regarding whether or not the cluster should be labelled malicious. Completely unsupervised algorithms adhere to a common rule, positing that the most significant clusters are the only consistent ones. As a result, routine network occurrences exhibit no signs of an attack. In reality, individuals must determine which combinations should be deemed abnormal. Supervised machine-learning strategies require at least one learning stage to be completed before initiating the traffic model construction [9].

In most cases, learning concludes once internet traffic patterns are sufficiently processed. Numerous controlled anomaly-based systems for detecting network intrusions employ various machine-learning methods. The proposed approaches typically divide the attack detection process into two phases: the first, termed feature vector extraction, and the second, known as the method development phase. For instance, developers utilized information theory to identify cyber threats, employing a metric based on entropy and information gain. Outliers were identified using a linear classification algorithm [10].

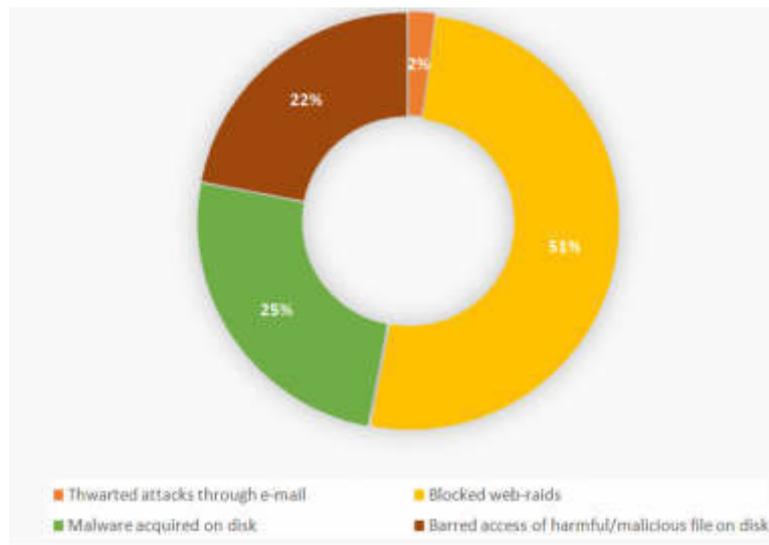


Fig. 1.1: The attack vectors worldwide in the year 2022

Unusual App activity is detected using a k-NN classifier with OS events, such as the quantity of started processes and system interactions. The authors applied a k-NN classification with metrics to the KDD sample to detect SYN Flooding, U2R, and remote-to-local (R2L) attacks. Classifiers and neural networks were jointly employed for spam detection. To identify DoS assaults, the researchers utilized a statistical approach. Feature vectors were used to train the Naive Bayes predictor, encompassing various UDP and TCP packets and their sizes. Attributes were successfully calculated using discrete wavelet transform (DWT) and combining retrieval based on a broad range of network properties. More sophisticated techniques, such as Hidden Markov Models, were used to identify denial-of-service and application layer assaults. Additionally, artificial neural networks are increasingly used in intrusion detection [11].

RBF artificial neural networks are applied to analyze network data for anomalies. Artificial neural networks prove effective in identifying UDP flooding incidents. Internet surveillance techniques based on support vector machines have also been investigated. The existing literature may offer solutions utilizing rough set theory and semi-supervised learning. The genetic method was employed to detect outliers. After the learning process, a set of rules is generated to characterize normal and aberrant traffic patterns in the network. The program underwent testing using data from the DARPA database. The feature vector encompassed the TCP connection's length, the amount of data transferred, and the link's originating and final IP addresses. The evolutionary algorithm and the correlation method were utilized to identify SQL Injection threats [12].

The subsequent sections of the article are organized as follows. The related works provide an overview of artificial intelligence, addressing supervised and unsupervised learning. The third section examines the proposed methodology for online safety, including network risk assessment and identifying and preventing infections. The fourth section explores the utilization of machine learning in cybercrime, with its simulated experiments and results.

**2. Related works.** Detecting intrusions and other abnormalities within the underlying Internet of Things (IoT) network is becoming an increasingly urgent concern. As IoT technology is deployed across various industries, the number of attacks and threats against it has also increased. Attack disruption attacks, data type searching, fraudulent authority, fraudulent functioning, scanning, espionage, and improper configuration are just a few of the attacks and irregularities that can compromise an IoT network. In this analysis, we compare the results of various machine learning models to assess their ability to predict incidents and anomalies on IoT devices reliably. Machine learning (ML) methods such as logistic reconstruction, SVM, RF, decision trees, and neural network training have all found applicability in this context. Performance levels are compared and

contrasted using criteria for precision, recall, reliability, F1 scores, and the area under the receiver operating characteristic curve [13].

Random Forest demonstrates superior performance across various metrics, even when the accuracy of the two approaches is comparable. In particular, the RF technique outperformed other methods' ability to detect and clear assaults effectively. The comprehensive study concluded that, based on the dataset analyzed, the RF approach is recommended for addressing intrusion issues in IoT networks, leading to these findings. It provided more accurate sample-level predictions for denial of service (DoS) and normal instances than any previous approach. These results suggest that RF analysis is well-suited for this investigation [14].

This study employs traditional ML methods on the dataset and then compares them without introducing any novel methodologies based on this dataset. Consequently, there is a need for further research to develop a reliable detection method, including a deeper exploration of the framework's general structure. The data from a simulated setting presents potential challenges when transitioning to real-time data applications. Addressing these challenges requires a more empirical study focusing on real-time data to provide a more comprehensive solution. The varying patterns of behaviour exhibited by different micro-services in the network contribute to the irregular behaviour of IoT services at different times. More research is necessary to understand these dynamics thoroughly. Despite achieving an impressive 99.4 per cent accuracy, the study's findings indicate that while RF performs well compared to other approaches, its continued effectiveness is not guaranteed with large datasets or unexpected scenarios. Therefore, additional research is suggested [15].

The hardware and computer software can be the foundation for an intrusion detection system. The primary objective is to monitor a system or a system of systems to ensure that no malicious or policy-breaking behaviour goes undetected. Intrusion detection systems (IDS) consider various network-related variables, such as source addresses, procedures, and banners, to assess the irregularity of behaviour. The fundamental goal of any security detection system is to achieve the highest level of precision while minimizing false alarms [16].

In intrusion detection, the study aims to identify the essential building blocks required for creating a detection framework. The model utilizes an ensemble method for identification that requires significantly less processing power, addressing challenges faced by conventional group-based surveillance algorithms. The article employs the Chi-square feature selection and a classification array consisting of SVM, modified naive Bayes (MNB), and LPBoost to construct an intrusion detection model. Chi-square feature selection analyzes variance to control the importance of each feature, focusing on those that matter in determining a class. Laboratory findings demonstrate the LPBoost ensemble's superior precision compared to baseline classifications. The most effective method for predicting the class label involves using a majority voting system, such as support vector machines, multi-network averages, or LPBoost, instead of a single classification. Given the substantial class disparity throughout the network traffic, this is preferable, providing a more robust approach than a single classification [17].

In the contemporary era marked by a digital revolution, knowledge storage, accessibility, and dissemination over the web have experienced significant and exponential growth. The emergence of innovations based on the IoT has further contributed to eliminating digital barriers and enhancing data and information exchange across pervasive networks. These IoT innovations have notably improved communication between devices, leading to heightened concerns among consumers regarding data theft, privacy breaches, and the secure transmission of their data and information over the web [18].

Various approaches, including deploying detection and prevention devices, are employed to combat data theft and other online data safety risks. This research specifically compares and contrasts two types of intrusion detection systems. The first type utilizes a machine learning strategy called SVM, while the second type employs an association rule data mining methodology known as Apriori. Both methods have originated from academic research. Research findings indicate that SVM surpasses Apriori in terms of accuracy, while Apriori demonstrates greater efficiency in testing duration [19].

Over the past decade, there has been a significant increase in individuals relying on the internet as their primary information resource. Technological advancements such as the IoT and high-performance computing (HPC) have accelerated the pace of online data access and the sheer volume of information generated. This surge in data production and accessibility has made safeguarding sensitive information from unauthorized access more challenging. Consequently, developing analytics and innovative models to mitigate threats posed by attackers

and scammers to private data stored online has become imperative [20].

The approach to enhance security is to limit the quantity of data stored online. In this study, the authors evaluated and compared the performance of two feature and two categorization methods for intrusion detection, aiming to mitigate the impact of data theft and other real-time network security breaches on computer systems. Leveraging the gathered information, researchers developed a robust intrusion model to prevent real-world attempts at network penetration in live software settings. Cyber-physical systems (CPS) integrate digital information processing capabilities and communication networks with tangible components and operations. Among the various anomalies that can impact the proper functioning of these systems, safety breaches and breakdowns are particularly common. While extensive research has focused on defect diagnosis and security analysis in CPS independently, the crucial and timely challenge of distinguishing between different sources of irregularities, such as flaws and attacks, remains inadequately addressed [21].

This study concentrates on the energy-aware smart home (EASH) system and its internal communication ecosystem. Specifically, the authors define the challenge of distinguishing between component breakdowns and network attacks in EASH, considering how these elements influence information exchange patterns. Key contributions of this paper include a formal demonstration of the relationship between these irregular sources and an ML-based methodology for identifying the issue. The developed system undergoes testing in both simulated and real-world laboratory environments, with findings indicating the potential to achieve a classification accuracy rate exceeding 85% [22].

Analysis of the classes and features employed in the proposed method suggests significant classification precision improvements based on our laboratory findings. The results from these tests substantiate these hypotheses. The study aims to explain the outcomes of a task to differentiate between anomalies affecting an EASH system. This suggests that observed anomalies could stem from malfunctioning hardware or malicious network activity. A discrimination operator was developed after studying the connection between various abnormalities and their impacts on the system's information channels [23].

Utilizing ML-based categorization algorithms sets the approach apart. This research concludes that supervised machine learning methods offer a viable option for distinguishing between faulty and attack classes with high reliability. Explorations into misclassifications of cases with similar effects on the network were conducted in simulation and real-time testbed experimental environments, encompassing analyses of aberrant classes and the considered properties. The findings suggest modifying description datasets by adding or removing features could enhance categorization outcomes [24].

The widespread adoption of IoT-connected devices is rapidly approaching, with IoT services expanding to achieve ubiquitous integration. The increasing popularity of IoT-enabled gadgets and services has led to many risks and attacks against them. While intrusions on the IoT are not new, the growing embedding of IoT in our lives and society underscores the critical need to enhance cybersecurity measures. Understanding potential threats and attacks against IoT infrastructure has become imperative. The research aims to explore the diverse dangers that could impact IoT devices and offerings, analyzing and characterizing the incursions and attacks directed at them. Recognizing potential threats is a prerequisite for implementing effective protective measures for the IoT. This article focuses on cybersecurity threats to the IoT, with the primary objective being identifying resources and inventory of potential threats, vulnerabilities, and attacks against the IoT [25].

A presentation highlighted security vulnerabilities associated with IoT devices and services, summarizing the most pressing issues affecting IoT safety. Secrecy, anonymity, and trust in entities were identified as challenging security aspects. The study emphasized that addressing privacy and security concerns is crucial for ensuring the security and user-friendliness of IoT devices and services. The article also investigated cyber threats driven by the unique characteristics of cyberspace, including individuals, motives, and capabilities. Evidence suggests that threats from government intelligence services and organized criminal groups pose more formidable challenges than lone hackers, as they may target fewer predicted goals, with a single attack expected to have a more significant impact [26].

The study concluded that manufacturers and consumers are crucial in ensuring IoT security. To address the weaknesses in the current security infrastructure, new standards for the IoT are essential. This research aims to contribute to future endeavours to enhance the understanding of vulnerabilities in IoT infrastructure and assess the likelihood of attacks against IoT and the consequences of such assaults. Early considerations in

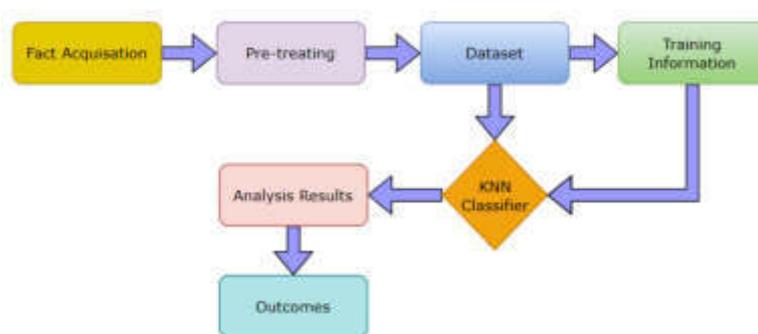


Fig. 3.1: The workflow of the proposed technique

the product creation process should include security measures such as access control, authentication, identity management, and a flexible reputation management system. This survey is anticipated to benefit security experts as it illuminates the most critical issues related to IoT security, offering a deeper understanding of the threats and their features [27].

**3. Proposed Methodology.** In this investigation, we employ an approach to intrusion detection and propose measures to enhance the security of innovative environments. The diagram illustrates the utilization of a K-nearest neighbours (KNN) classifier in the proposed method for intrusion identification. Following the dataset selection, a preliminary processing phase is implemented, converting certain symbols to numerical values to enhance the accuracy of the detection rate. The subsequent training phase involves the application of the KNN technique, training the model using 65% of the dataset. The final step involves evaluation, wherein the remaining 35% of the data assesses the model's accuracy. To further measure the performance of the proposed system, metrics such as the accuracy detection rate and the occurrence of four types of warnings are estimated in this phase. The architecture of the proposed model is explained in Figure 3.1.

Following the collection of information, preparation becomes necessary due to the unstructured nature of the features, which consist of text or numeric values. K-nearest neighbours (KNN) specifically deal with numerical data during the training and testing. Therefore, the objective of the preliminary processing stage is to convert symbols and characters into their numerical equivalents. The training and evaluation stages are crucial in developing an intelligent model. During the training phase, KNN enhances reliability, reduces false alarms, and assesses the false alarm rate. The dataset is then partitioned into a training set (comprising 60% of the total) and a testing set (representing 40%). Testing holds significance as it determines the model's effectiveness, with metrics such as precision, recall, and F1-score being computed at this stage.

**3.1. Dataset Description.** To establish an effective detection and prevention framework, having an accurate and current database is essential for concluding the operations of various cyberattacks. In this context, we leverage the multi-step cyberattack dataset (MSCAD), which comprises six individual files such as MSCAD.xlsx, N0, Scan, App01, App02, WB01, and WB02.

The annotated form of the provided dataset can be located in the MSCAD.xlsx file. Each of the six provided PCAP documents underwent analysis in OpenVPN. The timestamp is crucial in assessing whether network communication is malicious or benign. After processing these PCAP documents, the MSCAD dataset comprised 79 features with corresponding labels. The primary attack concept in MSCAD is rooted in a password-breaking model. MSCAD's secondary attack methodology predates the volume-based distributed DoS concept.

**3.2. K-Nearest Neighbors.** K-nearest neighbours (KNN) can effectively address regression and classification problems, offering simplicity in theory and practical implementation. However, its efficiency tends to decrease as the volume of data increases, resulting in sluggish performance. The training step, on the other hand, is highly time-efficient, as the data itself serves as a model for future identification, enabling spontaneous and randomized programming based on the provided information.

**3.3. Evaluation Metrics.** Assessment metrics are crucial in testing and refining classification algorithms within an intelligent intrusion detection system (IDS). These metrics are employed to evaluate the effectiveness of the IDS. In this study, precision is utilized to assess the effectiveness of KNN. Using Equation 3.1, the precision is defined as the percentage of detected activities correctly categorized [28].

$$Precision = \frac{\text{actions accurately categorised}}{\text{sum total of actions}} \quad (3.1)$$

The efficiency of a system can be quantified by calculating the occurrence of four distinct sorts of alarms (the ambiguity matrix). All four possible outcomes are true positive (TP), false positive (FP), true negative (TN), and false negative (FN) are identical.

$$TP = \frac{TP}{TP} + FN \quad (3.2)$$

$$TN = \frac{TN}{TN} + TP \quad (3.3)$$

$$FN = \frac{FN}{FN} + TP \quad (3.4)$$

$$FP = \frac{FP}{FP} + TN \quad (3.5)$$

In addition, Equation 3.6 and 3.7 are used to calculate accuracy and recall.

$$Accuracy = \frac{TP}{TP} + FN \quad (3.6)$$

$$Recall = \frac{TP}{TP} + FP \quad (3.7)$$

The F1-score is also used as a periodic mean of both recall and accuracy. The formula for the F1 score is as follows:

$$F1 - score = 2 \times (Precision \times \frac{Recall}{Precision} + Recall) \quad (3.8)$$

The relationship between the model size and recall, serving as a comprehensive guide that reveals the connections and dependencies within the framework, is represented in Figure 3.2.

The efficiency anomalies concentrating on the effects of employing distinct standard frameworks for each individual and URL response are shown in Figure 3.3. This figure is a pivotal reference point, shedding light on the complexities of using separate frameworks.

**4. Experimentation & Results.** This research divides the dataset into a training set and an evaluation set at a ratio of 56% to 39%. The proposed intrusion detection system is then assessed based on accuracy, precision, recall, and F1 score. The experiments are conducted on a computer with 64 GB of DDR4 memory, two gigabytes of dedicated graphics processing RAM, and two Intel Xeon 8 Core processors running at 2.43 GHz, housed in Dell PowerEdge T430 servers.

As indicated in Table 4.1, the model achieves an accuracy rate of 82.75%. This implies that the outcomes obtained using KNN machine-learning technique for detecting various attacks are generally reliable. To measure the system's capability in identifying intrusive situations, we utilize the MSCAD dataset, as demonstrated in the evaluation of the datasets. Figure 4.1 provides a comparison of our findings with the work of others.

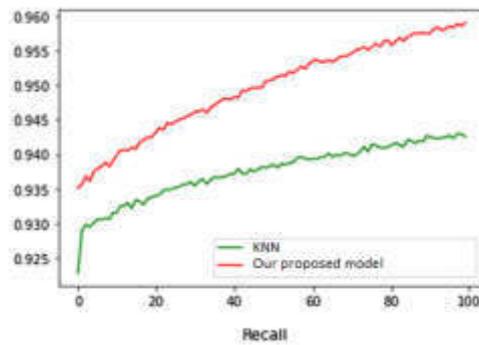


Fig. 3.2: The recall versus model size

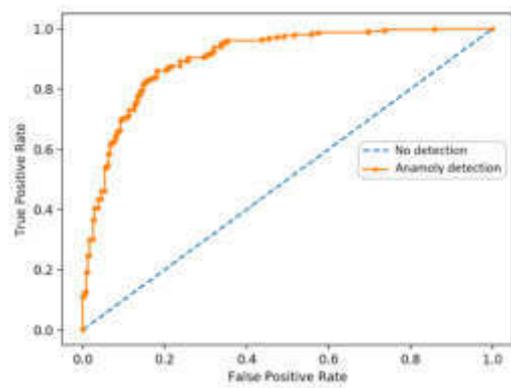


Fig. 3.3: Performance of finding anomalies using a separate normal model for every URL request

Figure 4.1 displays the accuracy values for three models: an IoT-based model with 94.6%, an ML-based model with 96.2%, and a proposed model with 98.5%. These accuracy percentages signify the proportion of accurate predictions made by each model, with the proposed model achieving the highest accuracy, followed by the ML-based model and the IoT-based model. The findings suggest that the proposed model performs exceptionally well, showcasing a strong ability to make precise predictions.

The model’s accuracy rate of 93.10% indicates the overall correctness of its predictions, showcasing the proportion of accurately classified instances in Table 4.2. The recall rate, at 90.22%, measures the model’s efficacy in correctly identifying relevant instances within the dataset. The F1-Score, reaching 91.42%, provides a balanced assessment by considering both precision and recall. Together, these metrics furnish valuable insights into the model’s multilayered performance, contributing to its predictive capabilities across diverse dimensions.

**5. Conclusion.** In this article, the application of AI is explored in security, considering both protective and threatening perspectives, and has discussed potential challenges faced by machine learning-based models. Machine learning offers an efficient approach to automating complex online operations for offensive and preventive purposes. As machine learning becomes an integral part of hackers’ cyber cache, it is anticipated that the complexity and diversity of AI-based attacks will increase. Computer education and security professionals are advised to stay updated on the modern developments in ML, including adversarial learning, to leverage this knowledge for national security. This study sets the stage for further research into the challenges of implementing scalable cybercrime systems using artificial intelligence in diverse operational contexts. This research focuses on detecting cybersecurity issues using a machine learning-based approach, specifically employing the KNN technique. The proposed system’s evaluation involves TP, TN, FP, FN, and measures such as F1-score, accuracy,

Table 4.1: Detection accuracy rates

Alarms style	Accuracy Rate (%)
TP	94.32
TN	91.97
FP	5.68
FN	8.03
Precision	82.75

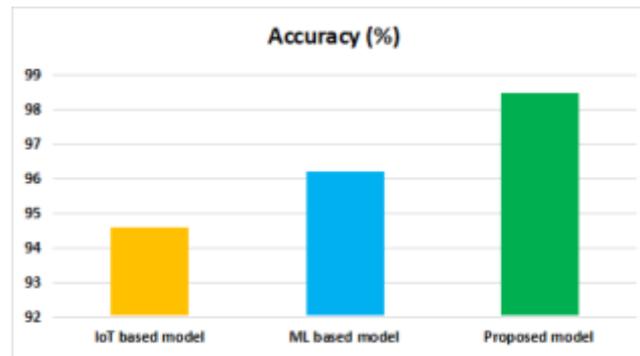


Fig. 4.1: Accuracy Rate Comparison

precision, and recall are calculated, revealing promising results.

**6. Acknowledgement.** The study was supported by Key R&D and Promotion Special Project (Science and Technology Research) in Henan Province (232102210146).

#### REFERENCES

- [1] B. Ahmad, W. Jian, and Z. Anwar Ali, "Role of Machine Learning and Data Mining in Internet Security: Standing State with Future Directions," *J. Comput. Networks Commun.*, vol. 2018, 2018, doi: 10.1155/2018/6383145.
- [2] Li, L., He, W., Xu, L., Ash, I., Anwar, M., & Yuan, X. (2019). Investigating the impact of cybersecurity policy awareness on employees' cybersecurity behavior. *International Journal of Information Management*, 45, 13-24.
- [3] Vocke, C., Constantinescu, C., & Popescu, D. (2019). Application potentials of artificial intelligence for the design of innovation processes. *Procedia CIRP*, 84, 810-813.
- [4] S. Dolev and S. Lodha, "Cyber Security Cryptography and Machine Learning ", In *Proceedings of the First International Conference, CSCML 2017, Beer-Sheva, Israel, June 29-30, 2017*.
- [5] Fenil, E., & Mohan Kumar, P. (2020). Survey on DDoS defense mechanisms. *Concurrency and Computation: Practice and Experience*, 32(4), e5114.
- [6] Vinayakumar, R., Alazab, M., Soman, K. P., Poornachandran, P., Al-Nemrat, A., & Venkatraman, S. (2019). Deep learning approach for intelligent intrusion detection system. *Ieee Access*, 7, 41525-41550.
- [7] Bagui, S., Kalaimannan, E., Bagui, S., Nandi, D., & Pinto, A. (2019). Using machine learning techniques to identify rare cyber-attacks on the UNSW-NB15 dataset. *Security and Privacy*, 2(6), e91.
- [8] AlErroud, A., & Alsmadi, I. (2017). Identifying cyber-attacks on software defined networks: An inference-based intrusion detection approach. *Journal of Network and Computer Applications*, 80, 152-164.
- [9] Aboueata, N., Alrasbi, S., Erbad, A., Kessler, A., & Bhamare, D. (2019, July). Supervised machine learning techniques for efficient network intrusion detection. In *2019 28th International Conference on Computer Communication and Networks (ICCCN)* (pp. 1-8). IEEE.
- [10] Gao, Y., Li, X., Peng, H., Fang, B., & Philip, S. Y. (2020). Hincti: A cyber threat intelligence modeling and identification system based on heterogeneous information network. *IEEE Transactions on Knowledge and Data Engineering*, 34(2), 708-722.
- [11] Seo, J. H., & Kim, Y. H. (2018). Machine-learning approach to optimize smote ratio in class imbalance dataset for intrusion detection. *Computational intelligence and neuroscience*, 2018.

Table 4.2: Performance Metrics for Model Evaluation

Estimation metrics	Rate (%)
Accuracy	93.10
Recall	90.22
F1-Score	91.42

- [12] Jemal, I., Cheikhrouhou, O., Hamam, H., & Mahfoudhi, A. (2020). Sql injection attack detection and prevention techniques using machine learning. *International Journal of Applied Engineering Research*, 15(6), 569-580.
- [13] Alsouda, Y., Pllana, S., & Kurti, A. (2019, May). Iot-based urban noise identification using machine learning: performance of SVM, KNN, bagging, and random forest. In *Proceedings of the international conference on omni-layer intelligent systems* (pp. 62-67).
- [14] R. Primartha and B. A. Tama, "Anomaly detection using random forest: A performance revisited," *Proc. 2017 Int. Conf. Data Softw. Eng. ICoDSE 2017*, vol. 2018-Janua, pp. 1–6, 2018, doi: 10.1109/ICODSE.2017.8285847.
- [15] M. Hasan, M. M. Islam, M. I. I. Zarif, and M. M. A. Hashem, "Attack and anomaly detection in IoT sites using machine learning approaches," *Internet of Things (Netherlands)*, vol. 7, p. 100059, 2019, doi: 10.1016/j.iot.2019.100059.
- [16] Y. Cheng, M. Naslund, G. Selander, and E. Fogelström, "Privacy in machine-to-machine communications A state-of-the-art survey," *International Conference on Communication Systems (ICCS)*, pp. 75-79, 2012.
- [17] Jiang, K., Wang, W., Wang, A., & Wu, H. (2020). Network intrusion detection combined hybrid sampling with deep hierarchical network. *IEEE access*, 8, 32464-32476.
- [18] Porkodi, R., & Bhuvanewari, V. (2014, March). The internet of things (IOT) applications and communication enabling technology standards: An overview. In *2014 International conference on intelligent computing applications* (pp. 324-329). IEEE.
- [19] De Almeida, M. B., de Pádua Braga, A., & Braga, J. P. (2000, November). SVM-KM: speeding SVMs learning with a priori cluster selection and k-means. In *Proceedings. Vol. 1. Sixth Brazilian Symposium on Neural Networks* (pp. 162-167). IEEE.
- [20] Nan, A. A., Shawky, M. M., Ahmed, A. M., & Ellaithy, D. M. (2021, December). Design and Implementation of High-Performance Computing Unit for Internet of Things (IoT) Applications. In *2021 International Conference on Microelectronics (ICM)* (pp. 258-261). IEEE.
- [21] Monostori, L., Kádár, B., Bauernhansl, T., Kondoh, S., Kumara, S., Reinhart, G., & Ueda, K. (2016). Cyber-physical systems in manufacturing. *Cirp Annals*, 65(2), 621-641.
- [22] Kabir, M. H., Hoque, M. R., & Yang, S. H. (2015). Development of a smart home context-aware application: A machine learning based approach. *International Journal of Smart Home*, 9(1), 217-226.
- [23] L. Zomlot, S. Chandran, D. Caragea, and X. Ou. Aiding intrusion analysis using machine learning. In *Machine Learning and Applications (ICMLA)*, 2013 12th International Conference on (Vol.2, pp. 40-47). IEEE, 2013, December.
- [24] Yadav, S. A., & Poongodi, T. (2021, April). A review of ml based fault detection algorithms in wsns. In *2021 2nd International Conference on Intelligent Engineering and Management (ICIEM)* (pp. 615-618). IEEE.
- [25] Tweneboah-Koduah, S., Skouby, K. E., & Tadayoni, R. (2017). Cyber security threats to IoT applications and service domains. *Wireless Personal Communications*, 95, 169-185.
- [26] Tewari, A., & Gupta, B. B. (2020). Security, privacy and trust of different layers in Internet-of-Things (IoTs) framework. *Future generation computer systems*, 108, 909-920.
- [27] Srivastava, A., Gupta, S., Quamara, M., Chaudhary, P., & Aski, V. J. (2020). Future IoT-enabled threats and vulnerabilities: State of the art, challenges, and future prospects. *International Journal of Communication Systems*, 33(12), e4443.
- [28] Itoo, F., Meenakshi, & Singh, S. (2021). Comparison and analysis of logistic regression, Naïve Bayes and KNN machine learning algorithms for credit card fraud detection. *International Journal of Information Technology*, 13, 1503-1511.

*Edited by:* Venkatesan C

*Special issue on:* Next Generation Pervasive Reconfigurable Computing for High Performance Real Time Applications

*Received:* Sep 26, 2023

*Accepted:* Dec 5, 2023



## SECURITY AND PRIVACY OF 6G WIRELESS COMMUNICATION USING FOG COMPUTING AND MULTI-ACCESS EDGE COMPUTING

TING XU\*, NING WANG †, QIAN PANG‡ AND XIQING ZHAO§

**Abstract.** The challenges surrounding the confidentiality of data transmission in the context of the upcoming sixth-generation (6G) wireless networks are proposed in this research. The study explores the potential role of blockchain systems in enhancing data security. It examines the integration of machine learning (ML) techniques to address the growing complexities of handling massive data volumes within the 6G environment. This research involves a comprehensive survey of existing strategies for maintaining data confidentiality in automotive communication systems. It further investigates an analysis of confidentiality approaches inspired by the 6G network architecture. The study examines the potential security implications of the Internet of Everything (IoE). It evaluates current research issues related to safeguarding data confidentiality within the framework of 6G communication among vehicles. The exploration involves reviewing ML techniques and their applicability in resolving the data processing challenges inherent in the 6G wireless network environment. The proposed work reveals the increasing complexity and variability of the 6G wireless network environment, leading to potential challenges in protecting private and confidential data during communication. It highlights the promising role of blockchain systems in addressing data security concerns within the 6G network context. Additionally, the study underscores the transformative potential of integrating ML techniques to handle the massive data volumes generated within the 6G ecosystem. The research highlights the importance of these technologies in mitigating data security risks and ensuring the confidentiality of information exchanged within the 6G communication framework.

**Key words:** Multi-access Edge Technology, Safety and Security, Internet of Everything (IoE), Wireless Communications, Cloud Computing Environment

**1. Introduction.** The extensive availability of wireless communication devices has had a profound impact on individuals' daily lives and has accelerated the expansion of the manufacturing sector. Businesses engaged in online communication and entertainment have particularly gained the rewards of the current fourth-generation (4G) infrastructure. The upcoming fifth-generation (5G) communication networks will order the Internet of Things (IoT), autonomous vehicles, and both virtual and augmented reality (AR/VR) applications. However, the absence of a personal touch in 5G communications has urged researchers and companies to anticipate the forthcoming phase of wireless communication. Despite its numerous benefits, 5G technology lacks a human element [1, 2].

The connectivity framework of the 6G generation transforms "linked objects" into "associated intelligence." Substantial investments in research and development for the 6G communication system are primarily directed towards the press, the product, and the fundamental building blocks. The emergence of advanced Intelligent Computing, the capability to gather, transmit, and assess information, and the facilitation of a diverse array of applications and intelligent services are all anticipated advancements with the introduction of 6G technologies. Human advancement will be ordered in the 6G paradigm above information, technology, and services. It is anticipated that the 6G communication infrastructure will strengthen confidentiality and security by eliminating any vulnerabilities associated with these issues, thus enhancing the overall reliability of the network. However, if conventional approaches to machine learning are used, the central server structure could be susceptible to privacy and security risks originating from various internal nodes, potentially resulting in a single point of failure [3].

Traditional artificial intelligence methods are inadequate for processing vast amounts of accumulated data.

---

\*Hebei North University, College of Information Science and Engineering, Zhangjiakou, Hebei, 075000, China

†Zhangjiakou Branch of China Mobile Communications Group Hebei Co. Ltd, Zhangjiakou, Hebei, 075000, China

‡Hebei North University, College of Information Science and Engineering, Zhangjiakou, Hebei, 075000, China

§Hebei North University, College of Information Science and Engineering, Zhangjiakou, Hebei, 075000, China (Corresponding author, [xiqingzhao63@163.com](mailto:xiqingzhao63@163.com))

Consequently, the sole solution presented by 6G involves the implementation of distributed artificial intelligence, with all customer-sensitive data stored within a set of instructional equipment close to the vehicle. The upcoming "sixth generation (6G)" of mobile phone networks, the successor to 5G mobile communications, represents a significant departure from the current technological landscape. 6G is anticipated to introduce cost efficiency, heightened security, and enhanced privacy to the system. Achieving this objective necessitates the integration of advanced technologies within the wireless network infrastructure. The ongoing deployment of automated, self-driving vehicles equipped with navigation systems, voice recognition structures, multiplayer games or movies, shopping capabilities, advertising functions, surveillance systems, and weather forecasting systems aims to reduce accidents, fatalities, and property damage. In today's increasingly automated environments, machine learning, often combined with artificial intelligence (AI), is a recent innovation in this domain [4].

Deep training is a form of artificial intelligence capable of discerning patterns and characteristics using vast datasets generated by lens detectors, ultrasonic detectors, and other devices. Automated driving, facilitated by autonomous vehicles (AVs), enables real-time adjustments of an automobile's acceleration and braking capabilities based on road conditions. The accuracy of judgments relies on both the input information and the capabilities of the deep learning (DL) model. Over the past few years, the focus on interconnected vehicles within the internet-of-vehicle (IoV) program has shifted towards integrated cognition. Data exchange and communication between mobile nodes will soon incorporate a hybrid approach, combining vehicle-to-anything (V2X) and designated limited-range transmission. Wireless access in vehicular environments (WAVE) relies on the 802.11p IEEE protocol, which will require updates to manage the substantial data flows across the internet. Anticipated advancements in 6G technology are expected to enhance V2X capabilities, highlighting privacy-protective features in these scenarios [5].

The beginning of self-driving vehicles represents an increasing technology that, upon full integration, promises customers a seamless and stress-free travel experience courtesy of machine learning and multi-access edge computing (MEC). By feeding back collected data to edge structures, MEC facilitates the automation of vehicles through edge intelligence (EI), enabling them to anticipate and respond appropriately. The implementation of EI technology has the potential to strengthen the network's dependability and generally reduce latency. However, despite the manifold benefits outlined, the pursuit of establishing an automated 6G ecosystem faces various challenges. Ensuring the reliability of wireless networks, preserving the privacy of individual nodes, safeguarding sensitive data, and other related factors necessitate careful consideration [6].

A relatively young AI technology has swiftly emerged as one of the most powerful tools in its domain. Its adaptability has led to its widespread use in diverse fields, such as wireless network security, traffic monitoring, and dynamic topology alert systems, as well as for ensuring data reliability and privacy. These applications represent only a fraction of AI's potential. While initially intended for monitoring geographically dispersed networks, AI is now commonly employed to improve networks' overall efficiency and functionality. The network's initial structure was deliberately kept simple to enable upgrades at any node, anytime. Since integrating AI, the network has been continuously monitored and fortified against potential threats. Encryption methods are frequently employed to detect and prevent suspicious activities within wireless networks, including monitoring network load and traffic volume to identify and counteract denial of service (DoS) attacks. AI-based technology can complement existing network security strategies, ultimately reducing wireless security risks [7].

Artificial intelligence, machine learning, and deep learning algorithms, previously unattainable services, have become feasible. These innovative techniques have enhanced our understanding of wireless environments and enabled accurate predictions across various activities. Tasks such as capacity classification, resource management, and data security stand to benefit significantly from the application of such technology. Implementing these methodologies can positively impact the quality of experience (QoE) and enhance network reliability. Furthermore, advancements in machine learning have automated the creation of communication models, facilitating the development of increasingly sophisticated methods for connectivity. Additionally, users can now request internet access, prompting the system to identify relevant devices and service-oriented applications and provide them to the user in an orderly fashion. Previously, users had to expend significant time and effort conveying their requirements to network technicians for tailored system maintenance [8].

The aim of achieving greater adaptability and swiftness in this sphere is realized by implementing a Software-Defined Networking architecture. By enabling easier access to complete permissions, access providers

can better cater to the needs of clients with stringent requirements, which stands as the primary objective of this initiative. By integrating multiple encryption mechanisms, SDN and AI can streamline various operations, including load distribution, process management, and network security against potential threats. Some estimates suggest that over 55 per cent of access providers plan to incorporate AI into their infrastructures to streamline current processes. The recent spread of AI-powered devices with embedded connectivity has sparked notable advancements in in-vehicle connectivity. This modern technology has significantly enhanced the performance of various blockchain-powered devices, including IoVs and software-defined networks (SDNs). While the current 5G technology adequately supports existing devices, it may face challenges accommodating future technologies such as 3D video chats. The convergence of AI and IoT has led to a wider array of interconnected devices capable of participating in and benefiting from online communication [9].

Hence, in the future, 5G will only be able to support a limited data volume or a large customer base. Under the current 5G standard, administrators can manually configure and optimize their networks. Consequently, manually managing a sprawling, dynamic, and widely dispersed network poses significant challenges. In a scenario beyond 5G (B5G) or with the advent of 6G, the integration of AI technology is anticipated to render these vulnerabilities obsolete.

**2. Related works.** In vehicle networks, numerous drivers can engage in real-time information exchange while providing a diverse range of conveniences to fellow motorists. These services include GPS navigation, ridesharing, valet parking, real-time traffic updates, and digital vehicle diagnostics. The fifth generation (5G) of radio communication technologies introduces significant enhancements in security, accessibility, and network capacity. Educational and corporate institutions have displayed considerable interest in 5G-supported vehicular networks (5GVNs), expecting them to have a transformative impact on the transportation landscape and foster various novel connectivity options. Simultaneously, advancements in sensor technology and the expansion of regional data collection systems have made gathering extensive information from vehicle users, including identity verification, status updates, and location tracking, an unavoidable reality. Regrettably, 5GVN still grapples with diverse privacy threats, leaving users' information, viewpoints, locations, and movements susceptible to potential breaches. The latest research findings examine 5GVN's architecture, features, and capabilities. It subsequently delves into the privacy objectives of 5GVN and the associated security risks, providing an in-depth analysis of existing solutions to preserve user privacy. Finally, potential avenues for further research are outlined to stimulate greater interest in this innovative design and its privacy concerns while encouraging grassroots initiatives to address these issues effectively [10].

The 5GVN represents a groundbreaking framework poised to revolutionize the scale of service deployment and vehicle communication. As 5G networks expand and electric vehicles become increasingly commonplace, 5G will become more deeply ingrained in our daily lives. Nevertheless, privacy risks continue to raise a spectrum of concerns for automotive users. This research extensively evaluates numerous privacy-preserving options for 5GVN, encompassing an exploration of its architecture, features, and capabilities. The study emphasizes the imperative of protecting user privacy, assessing the threats within 5GVN, and examining the problems and solutions ensuring user anonymity within the 5GVN context. Ultimately, the study highlights the need for continued attention and effort to address 5GVN and its associated security challenges [11].

The 6G represents cutting-edge technology with promising capabilities for addressing the demanding communication requirements of autonomous vehicles. Its potential benefits include minimized latency, reduced communication costs, and a strong focus on safeguarding user data in the dynamic 6G ecosystem. Various technologies, such as machine learning, deep learning, artificial intelligence, and blockchain, can be used to address these challenges. However, the increasing number of vehicles on the road and interconnected IoT devices has made ensuring a reliable flow of information over the Internet increasingly complex, particularly with the application of conventional ML techniques.

The authors examine the potential privacy implications of IoE within the context of 6G automotive interactions, highlighting ongoing research concerns in privacy preservation. Safeguarding consumers' privacy and security remains a principal consideration in wireless networks. While introducing 5G technology brought new threats to user safety and privacy, the advent of 6G network technology is expected to mitigate these issues. As a result, advancing an efficient and reliable identification mechanism becomes imperative in 6G-powered transportation communication [12].

The IoV introduces a novel computing model offering several advantages to drivers and users, including route suggestions, reduced traffic delays, and enhanced driving experiences. However, sharing IoV participants' data with external entities exposes them to potential privacy breaches. This article conducts a comprehensive analysis of the privacy risks associated with IoV. It proposes strategies for addressing these concerns, examining the privacy challenges of IoV's architecture. In this exploration of the IoV, the authors focus on ensuring the security of sensitive information. Initially, researchers investigate the structural intricacies of IoV, highlighting the inherent privacy challenges. The authors established a cohesive framework for safeguarding personal data, encompassing encryption, anonymity, and perturbation as essential privacy measures. Additionally, potential avenues are thoroughly examined for research, emphasizing IoV applications and advanced methodologies while prioritizing individual privacy requirements [13].

The paper addresses the privacy challenges associated with V2X communication in the cloud by introducing the security evaluation methodology with understanding considerations. Leveraging artificial intelligence on previously collected data enables the assessment of cloud-based uncertainties based on prior evaluations. Moreover, PAU accounts for privacy-preserving capabilities derived from moment-to-moment vehicle interactions. It presents a security aggregation technique that combines online and offline perspectives to enhance the precision of the security assessment. In these models, a method is established for generating the mixed zone by selecting nodes with a heightened awareness of confidentiality. The outcomes of our research validate the efficacy of our approach on various fronts, particularly in stimulating the system against malicious behaviour and slander-based attacks using historical data stored on the public Internet.

The rapid advancement of technology and modern vehicles underscores the critical need to ensure the security and reliability of the IoV. This environment inherently faces ongoing security challenges due to the vulnerability of wireless networks, leading to concerns about personal safety, protection, and anonymity on roadways. Over the years, various privacy- and security-preserving techniques have been proposed to mitigate destructive player attacks, complex computations, and high communication costs, aiming to overcome the limitations of existing solutions [14].

Several approaches have been suggested by computer security experts, authors, and practitioners to address the security and privacy concerns [15]:

- (i) The proposed system integrates a structure akin to fog cloud vehicular ad hoc systems.
- (ii) The solution incorporates an identity protection mechanism based on public key cryptography for securing user information.
- (iii) To assess and rank four distinct algorithms for dynamic routing using a simulator, tracking the findings for each protocol and conducting a graphical analysis of the collected data throughout the simulation. The combined impact of ECC deployment and the optimal routing protocol on VANET performance are evaluated in this context.

The VANETs safeguarding many prior initiatives have merely repurposed conventional security methods that have proven effective without accommodating the network's specific nature. Several solutions advocated by security specialists rely on trusted platform modules (TPMs), perceived as single points of failure, or involve intricate computations impacting VANET efficiency. In response to our examination of VANET architecture and its associated security challenges, an approach is formulated that fulfils the requirements for confidentiality while ensuring the protection of individual user privacy [16, 17].

A novel system, called the fog cloud, has been introduced to incorporate fog computing. This was accomplished through experiments, inspections, and evaluations in Network Simulator 3 to identify the optimal routing protocol, addressing the intricacies of security solutions. The dynamic source routing (DSR) algorithm was determined to be the most effective approach when implementing ECC within the system. Leveraging encryption to uphold user confidentiality and ensure anonymity is based on insights from the study on each of the employed methodologies [18].

### 3. Proposed Methodology.

**3.1. Internet-of-everything (IOE) Security Concerns in the Fifth Industrial Revolution.** Industry 4.0 has seen a surge in standards owing to technological advancements such as artificial intelligence (AI), the IoT, and cloud-based computing. Its core aim is to make various businesses "smarter," overseeing an array of tools and technologies throughout their life cycles. Automation is the key to reducing the need for human

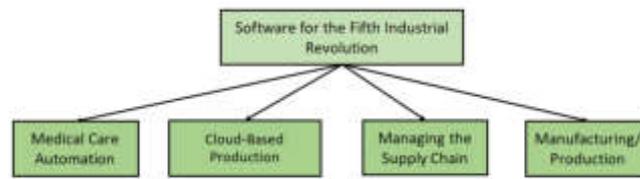


Fig. 3.1: Applications of Industry 5.0

involvement in labour-intensive processes. The goal is to leverage AI to bridge the performance gap between computers and machine learning (ML) devices, ultimately achieving higher overall efficiency. Looking ahead to Industry 5.0, the focus is on integrating human and artificial intelligence. Many experts anticipate that human interaction will be pivotal in this next industrial revolution. Industry 5.0 is self-assured to boost production, catering to human and robotic adaptability. This will facilitate seamless information and feedback exchange between humans and machines, ultimately enhancing productivity. Within the Industry 5.0 framework, emphasis is placed on optimizing the quality of final products by differentiating tasks that require research and innovation from mundane, time-consuming activities. Skilled labour is encouraged, as people will be responsible for coding instructions for robots, promoting mass customization in manufacturing. Unlike the previous phase, Industry 4.0, which focused on mass production, Industry 5.0 significantly focuses on personalized and individualized end products, as depicted in Figure 3.1.

Machine learning (ML) is an increasingly reliable and prevalent innovation in the medical field. ML-based models are now used to diagnose patients, leading to faster and more accurate assessments. However, achieving a fully automated environment requires more than just these models. The concept of Industry 5.0 opens the possibility of automated processes, including patient examination and medication administration, based on the findings. In recent years, high-tech devices equipped with sensors have been developed to monitor patients' conditions, such as smartwatches and fitness trackers. These devices provide valuable data that aids physicians in conducting thorough examinations. In critical situations, these devices can communicate with each other and alert the medical team. According to Industry 5.0, robots could independently perform crucial surgical procedures, enabling human physicians to focus on innovation and research.

Internet manufacturing is a new technology that integrates IoT devices and cloud-based approaches to provide an automated virtualization experience to end-users. This collaboration among international stakeholders aims to reduce production costs and increase overall productivity. Cloud computing in content creation offers various advantages, including improved quality, reduced production expenses, enhanced security, and reliability. It also helps minimize environmental impact by preventing the mass production of defective items. Moreover, integrating human insights into IoT and artificial intelligence development can help address issues associated with unexpected mishaps. By combining human intelligence with machine intelligence, many unforeseen disasters, including earthquakes, tsunamis, and other natural calamities, can be efficiently mitigated.

**3.2. New 6G methods for protecting security.** The extensive wireless connections within future 6G environments, particularly in various vehicles, will raise the bar for ensuring privacy and security. The forthcoming 6G networks will require specific features to address potential security breaches. While several advanced techniques have demonstrated their capabilities within the 6G context, their ability to uphold security and privacy remains uncertain. Conversely, technologies such as cloud computing and system software are under scrutiny concerning the privacy aspects of 5G.

In summary, the potential solution for mitigating the reliability challenge in the upcoming 6G environment involves addressing key concerns, including reducing privacy risks by implementing available and contemporary solutions. Figure 3.2 provides insights into how certain common technologies could minimize privacy and security risks.

The Blockchain approach has experienced significant growth in distributed database technology, gaining traction in the mobile phone industry through various publicly available DLTs. Leveraging blockchain to

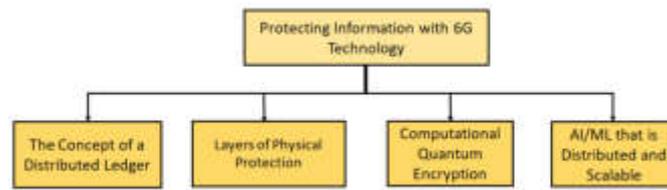


Fig. 3.2: Secrecy Tools for the 6th Generation

integrate multiple services into a 6G network ecosystem while maintaining security offers numerous advantages. However, incorporating machine learning (ML) or additional data analysis techniques might lead to potential vulnerabilities in 6G networks despite the benefits of integrating AI approaches. Several machine learning methods are susceptible to various forms of attack during both the training and testing phases. Consequently, verifying the authenticity and source of data before subjecting it to various artificial intelligence methods is crucial. The use of immutable data in DLT can help ensure data integrity, with this assurance being shared among the numerous parties involved in establishing the trustworthiness of AI robots in a dynamic environment. Drawbacks arise from users' exposure to a wide spectrum of security issues when employing DLT in a 6G context, potentially impacting the efficiency of the 6G ecosystem. Vulnerabilities in programs, limitations in programming languages, and numerous security and privacy ambiguities in internet connections are primary sources of these attacks. Such confusion can lead to compromised trust, financial losses in cryptocurrency transactions, or the unavailability of an internet connection or website.

Integrating digital currency within the 6G network architecture framework offers solutions for specific challenges and new opportunities. Public ledgers are more relevant than private ones regarding security routing concerns. For example, validating whether smart contracts have been updated becomes highly challenging since every node in a blockchain-based network is responsible for their validation. Smart contracts are an integral part of blockchain infrastructure, crucial for enabling automation and ensuring the legitimacy of transactions. Various reliability testing methods are used to search semantic faults to ensure the correctness of a smart contract. Moreover, rigorous monitoring of artificial intelligence (AI) blockchain nodes or malicious software with strict control over access and verification is essential. These solutions can withstand a wide array of attacks. Blockchain innovation enables the implementation of several security measures, including privacy by design and trusted execution environment (TEE), making it possible to construct more resilient networks. Below are some examples of networking types that a distributed ledger can accommodate.

There are various blockchain types, including open, closed, connection, and mixed networks. Utilizing different types of internet infrastructure gives rise to different privacy concerns. If, for instance, more than 50% of attacks against a particular 6G enterprise are executed through publicly accessible blockchains, proactive measures may become necessary. The subsequent step involves establishing either a consortium blockchain or a private ledger, which can enhance efficiency, reduce vulnerability to attacks, and better safeguard the confidentiality of a given dynamic node. Hence, introducing any blockchain system impacts specific risks and the diverse services offered in a 6G environment.

The emergence of quantum computing technology is anticipated to eventually replace all existing computational methods and cryptographic systems. It is suggested that vulnerabilities in 6G wireless networks could be detected, mitigated, and prevented using this approach. Networks leveraging 6G are proposed to achieve unprecedented dependability with quantum technology, enabling the transition from silent communication modes to cognitive-communication pathways. Several qualitative methods have been proposed to address the continual exponential issue in public key cryptography. Moreover, quantum computers are expected to strengthen their security infrastructure shortly. Based on its foundation in the atomic structure of information, quantum computing is widely expected to significantly anticipate novelty and security, enhancing the dissemination capabilities of 6G communication, which rests on quantum data.

Oblivious transfer is a conventional information dissemination method wherein the sender is unaware of the

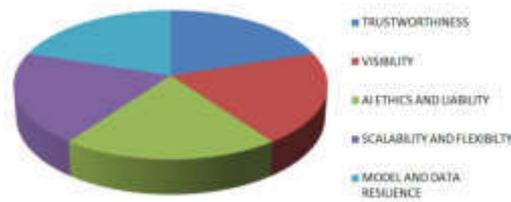


Fig. 3.3: Difficulties in applying ML/AI to 6G

specific information being transmitted. Yet, such information exchanges are not feasible in quantum computing, as any information leakage may disrupt a two-way connection. According to the quantum principle, systems lack a replicating property, rendering the precise replication of their states physically impossible. Thus, to successfully tamper with the data, an adversary must first extract its random quantum state and then replicate it into a duplicate without altering the original state of the data. A quantum collision could also occur if different values fed into a hash function yield the same result.

Numerous industries and academic institutions are investigating and creating Quantum attack solutions to address the increasing challenges of upcoming 6G communications. Several diverse quantum-proof algorithms are founded on lattice-based structures, hash values, or multiple variables. The issue related to computing on a lattice can be resolved by leveraging IoT devices, which offer enhanced performance, connectivity, and more. Since traditional random oracle models will not suffice with quantum-proof algorithms, it is crucial to authenticate security at multiple levels within a quantum-obtainable random oracle framework. This prevents adversaries from potentially querying the random oracle in non-relativistic time if the framework is unprotected.

The forthcoming 6G ecosystem will enable a wide array of automated tasks to be executed without human intervention, encompassing monitoring, computation, recovery, and optimization. Zero-touch infrastructure & services management (ZSM) is an emerging architecture that establishes fully autonomous wireless connectivity emphasizing cybersecurity through artificial intelligence and machine learning. As extensive sixth-generation mobile networks will generate massive volumes of data, wireless networks must incorporate artificial intelligence and machine learning to handle the resulting data overflow. The authentication checks in 6G connections can be enhanced using a diverse array of distributed cryptographic techniques based on artificial intelligence and machine learning. The primary benefits of this deployment approach are the precision and foresight of machine learning and artificial intelligence algorithms for cryptography within a 6G context.

There are challenges to overcome and rewards to gain when bringing AI and ML into a 6G setting. Some of the difficulties are described in applying AI/ML in 6G is shown in Figure 3.3.

The most demanding question that must be answered before using AI and ML in a 6G setting is whether or not we can have faith in these innovations to reliably and securely run a susceptible network.

**3.3. Analysis of privacy-protecting methods for 6G-powered automobile communication.** The global demand for cutting-edge autonomous vehicles equipped with advanced features has increased. The rapid advancement of virtual reality technology has raised concerns regarding security and privacy, thereby underscoring the importance of safeguarding both drivers' safety and privacy within a fully automated wireless environment. The extensive data collection from multiple devices to provide diverse consumer services necessitates implementing various security approaches that effectively protect this information without compromising user privacy. In the context of 6G networks, the transmission and management of voluminous data become significantly more complex, given the exponential increase in the number of end nodes compared to 5G networks. Balancing the delicate data security and privacy concerns within the 6G landscape while maintaining the current encryption techniques' privacy standards requires careful consideration of the data processing costs and the need for privacy.

Monitoring each device within the 6G wireless network poses a substantial challenge, potentially diminishing the efficacy and transparency of the data acquisition process. Consequently, adopting big data technologies

across distributed networks with a focus on ensuring privacy and security could increase the system's overall computation and transmission costs. Key obstacles to maintaining user security and privacy in a 6G network include the introduction of security vulnerabilities that may lead to the theft of users' personal information during extensive wireless communications. Additionally, the significant data accumulation facilitated by 6G technology could potentially compromise the privacy of individual nodes within the network as 6G ushers in a new era of AI-powered, high-tech devices, the deployment of various applications across the network's framework necessitates the development of innovative lightweight security measures at the network edge. Achieving this balance will require a stable environment prioritizing exceptional service provision while ensuring user safety and data privacy.

The careful management and restriction of access to data collected by various service-based applications is crucial for ensuring that these applications function as intended, especially concerning information related to the user's identity and immediate surroundings. In a 6G environment, blockchain technology presents a novel solution to address customer privacy concerns. While integrating blockchain technology into 6G networks offers numerous benefits, there are also potential drawbacks. Blockchain can potentially serve as a safeguard in maintaining data security within a 6G landscape. Blockchain enables users to maintain security without compromising their physical location or identity by functioning as a unified messaging system across the network.

It is important to note that blockchain, a form of distributed ledger technology (DLT), operates on a publicly accessible platform. Consequently, all user data collected within this framework may become publicly available. While 6G technology strives to establish a more secure architecture that enables reliable service-based and network-based activities, there remains a significant risk to the security and privacy of the data collected and the relationships established between dynamic nodes. Despite the potential for digital currencies to facilitate privacy objectives among distributed service-oriented devices, potential security and privacy concerns persist.

Digital certificates, CoinJoin, and similar techniques can effectively alleviate privacy concerns for consumers. One crucial area where privacy issues have significant implications is ML approaches, supported by cutting-edge and rapidly expanding AI techniques in the 6G environment. When utilizing a flexible 6G system, it becomes possible to securely store the user's personal information in specific devices distributed through AI/ML methods. However, there remains a risk that AI and ML-based methods might be susceptible to attacks from other AI and ML methods, resulting in data loss, alteration, or other unintended consequences. Various forms of artificial intelligence can be employed to safeguard users' private data, including messages, geolocation information, and other sensitive data. Within an AI/ML-driven framework, an attacker typically seeks to predict the next unpredictable outcome generated by the knowledgeable model during the training or testing phases.

The concept of quantum computing can be personalised to enhance reliability and increase efficiency in a 6G setting. As researchers explore better and more robust ways to protect users' privacy, these diverse and heterogeneous approaches can help maintain 6G wireless networks over time. However, the anticipation of integral applications and vast scenarios within 6G communication networks has amplified the significance of privacy protection and highlighted notable privacy-related issues, such as anonymity, node connections, and the lack of reports on the visibility of changing nodes within the wireless infrastructure. As a result, many individuals remain cautious about providing their personal information online. Consequently, employing randomized patterns to provide consent for sensitive data sharing has become a common approach to privacy protection. Several proposed approaches address confidentiality-related issues to mitigate the potential of adversaries studying private information.

## 4. Experimentation & Results.

**4.1. Recommended model for asymmetrical privacy protection.** In this ecosystem, we present DPSSmartCity, an SDN-based technology that protects user confidentiality and provides network administrators more flexibility. Below are some code snippets to illustrate the core concept of our proposed solution. To address the limitations of current research, we introduce the DPSSmartCity approach, designed to establish a thriving and efficient smart city leveraging the Internet of Things (IoT). The IoT infrastructure within a smart city ensures the security of individuals' personal information. The DPSSmartCity framework comprises two main components:

- 1) A software-defined networking (SDN) enabled IoT-based smart city.

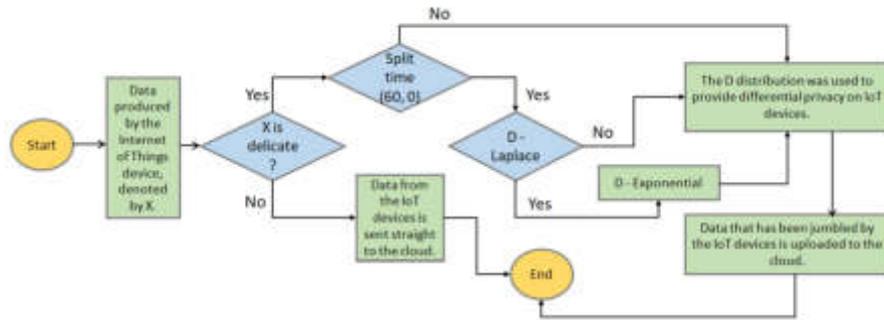


Fig. 4.1: Dynamics of Privacy-Preserving Mechanisms in Smart City IoT Infrastructure

2) A strategy for safeguarding individual privacy within the wired smart city, particularly when utilizing IoT devices with varying levels of confidentiality.

In the provided example, all IoT nodes are interconnected through an OpenFlow bridge, which is then linked to a central hub known as an SDN administrator. The software-defined networking (SDN) controller communicates with the cloud computing environment. IoT devices are interconnected in our smart city scenario, and their data is routed to OpenFlow switches through SDN. The SDN controller oversees and regulates these switches via wireless or wired connections. The correlation, highlighting the benefits of SDN infrastructure, can be connected within this innovative city concept, illustrated in Figure 4.1.

These advantages include enhanced mobility, remote administration, and centralized control while ensuring data remains confidential. The SDN controller maintains constant connectivity with both Cloud links, enabling it to receive instructions from the cloud and relay them to IoT devices, facilitating bidirectional data flow. First, the IoT-based smart city is integrated with the SDN paradigm, followed by the embedded approach that upholds user privacy, as depicted in the confidentiality technology flowchart presented in Figure 4.1. Upon the control system receiving modified data, it is directed to the SDN, eventually making its way to the cloud for further analysis by the SDN controller. Conversely, if the data isn't highly sensitive, the IoT device transmits it directly to the Internet.

When an IoT device generates private data, it employs differential privacy (DP) to safeguard this information. Within a connected urban environment, each IoT device has two options to choose from to protect user privacy:

- 1) A differentially private approach utilizing the Laplace probability principle.
- 2) A differentially private technique employs the inverse exponential probability.

The SDN controller updates the dissemination of the currently active differential privacy approach every hour. As a result, the distribution pattern of the utilized approach in vulnerable IoT devices fluctuates between Laplace and exponential probabilities on a minute-to-minute basis. Following applying one of these DP techniques, the device transmits the data directly to the Internet. Due to the continuous updates to the designated privacy protocol across all connected devices, the dynamic environment's fluctuations and intentional differential privacy intrusion are more effective in preventing the leakage of sensitive information. The detailed procedure can be expressed as follows.

Let us consider a scenario where information is transmitted between IoT devices and recipients as quickly as possible, but there are instances where information transmission may require several minutes. If data transfer exceeds a certain time threshold, the protocol switches to the alternative method during data transmission. Therefore, if the required time for communication delivery is denoted as  $\tau_i = \min p_i, M$ , where it represents the time needed for successful communication delivery, notably, the value of  $p_i$  is directly linked to  $M$ , indicating that as the value of  $M$  increases, the probability of the approach being unsuccessful rises. Consequently, the value of  $p_i$  approaches one as the value of  $M$  approaches its maximum. Additionally, we interpret the symbol  $T$  as signifying an extended period during which any protocols are vulnerable to breach, further increasing the complexity of theft attempts. The initial decision regarding the protocol at the onset of the attack diminishes

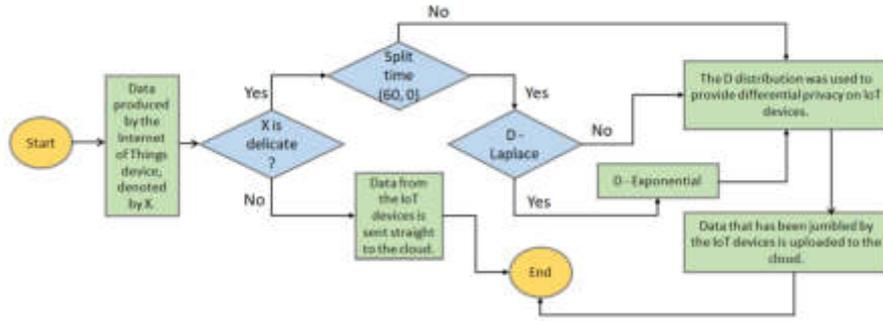


Fig. 4.2: Computational effort and T's relationship

Table 4.1: CPU Usage Trends for Different Device Configurations

Time (s)	Number of Devices	CPU Usage (%)
1	500	22
4	500	20
14	500	28
1	165	16.5
4	165	16.5
16	165	16.5

the chance of success to one-quarter. Therefore, if denotes the likelihood of success at any given time (regardless of the technique employed), the probability of success is a quarter of the anticipated value of each of the remaining four techniques. The escalation in processing costs directly results from the mounting operational load. Moreover, interpreting as the probability of the entire process failing, will oscillate between its minimum value for lower values of  $M$  and its maximum value of 1.0 when  $M$  equals the value of  $T$ . Employing the method across multiple modes will help determine the optimal scenario, balancing computational expense criteria and the risk of total operation failure, yielding the most favourable outcome.

**4.2. Simulation and Evaluation .** The limited computational capability of IoT devices is a widely recognized fact. The current predominant concern revolves around the associated costs, which can only be comprehensively understood through practical application. In this section, we investigate the evaluation of the impact of our approach on IoT devices by examining its overhead. Upon implementing the DP smart city method, there is a discernible increase in the workload, which can be measured as a proportion of the devices' aggregate median augmented CPU utilization. The fundamental attributes of the simulation and the corresponding expenses will be explained below.

The SDN concept in a cutting-edge urban environment is introduced, capitalizing on the highly dynamic nature of the smart city's setting. The proposed process is simulated in C# using Visual Studio 2019. Expanding the total count of IoT devices, we observed a resultant increase in throughput, as depicted in Figure 4.2.

The relationship between the overall probability and the value  $T$  and the associated overhead is illustrated in Figure 6. The average CPU usage across the entire network is displayed in Figure 4.3.

To maintain conciseness, Table 4.1 presents the data on CPU usage in just four different states, although the general trend remains consistent across multiple scenarios. Notably, the proportion of CPU time utilized during processing tends to vary slightly across different runs. For instance, during the first, fourth, and fourteenth seconds, 22%, 20%, and 28% of CPU time were consumed by 500 devices, respectively. The current operational costs for 165 devices are as follows: there was a 16.5% increase in periods 1, 4, and 16 (s), respectively. Both the pre-execution surge related to data retrieval and the post-execution surge related to recording are transient,

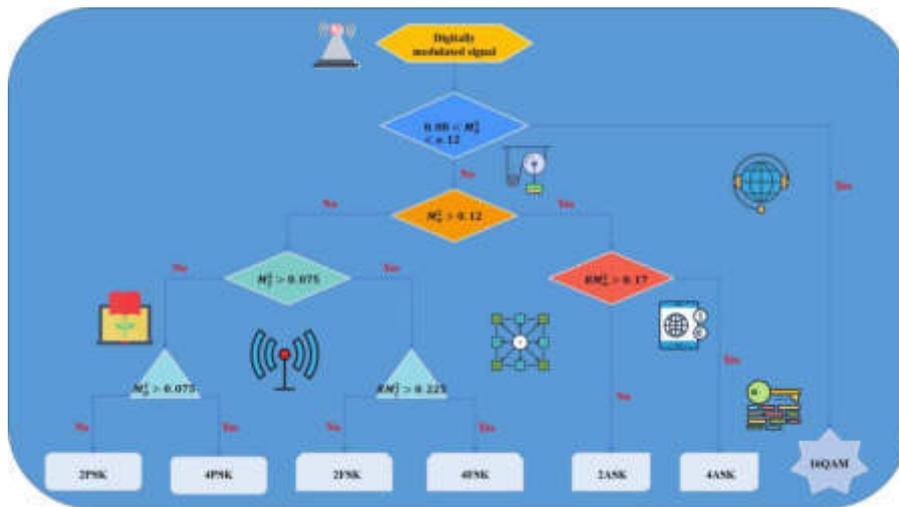


Fig. 4.3: Overall probability’s relationship to the value T

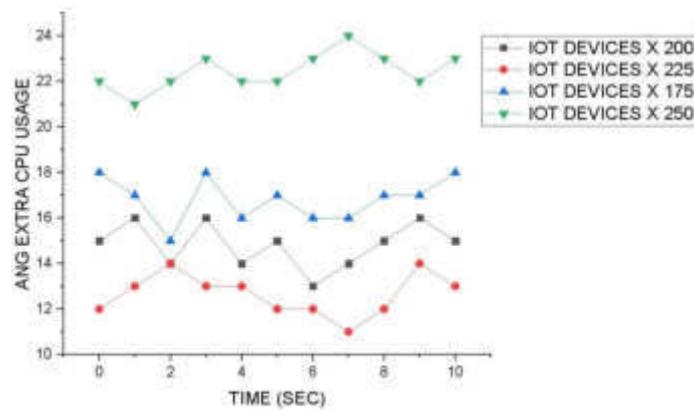


Fig. 4.4: Overhead analysis

with a slight potential capacity boost during this time. Our findings underscore the substantial cost implications for IoT devices.

As depicted in Figure 4.4, the proposed technique imposes only a 9–19% additional cost on IoT devices. Consequently, our approach can be implemented on IoT devices capable of accommodating the extra processing load. Moving forward, we aim to analyze the proposed approach from multiple perspectives, including assessing the proportion of successful hackers attempting to access confidential information. Additionally, we are keen on examining our approach’s performance in scenarios where the allocation time may need to be adjusted.

**5. Conclusion.** Ensuring the security and privacy of consumers has always been integral to maintaining a reliable internet connection. With the advent of 5G networks and the anticipated rollout of 6G networks between 2026 and 2030, global connectivity can expand dramatically, linking various digital and physical systems, including automated vehicles and sophisticated technologies. This expansion escalates the complexity of potential threats, ranging from learning-enabled attacks to significant data breaches. In the context of 5G networks, the primary challenges include facilitating high capacity, strong connectivity, low latency, robust security, minimal energy consumption, extensive knowledge integration, and reliable networking. This study sought to address safeguarding users’ privacy in 6G vehicular communication, considering historical perspectives

and future possibilities. It explored various advanced 6G innovations and their potential applications within the framework of Industry 5.0. Furthermore, the study conducted a comparative analysis of several privacy protection methodologies within the context of 6G mobile communication, examining the current state of the art and identifying the most effective approach for securing privacy in 6G-driven automobile communication settings.

## REFERENCES

- [1] Torres Vega, M., Liaskos, C., Abadal, S., Papapetrou, E., Jain, A., Mouhouche, B., ... & Famaey, J. (2020). Immersive interconnected virtual and augmented reality: A 5G and IoT perspective. *Journal of Network and Systems Management*, 28, 796-826.
- [2] M. Gheisari, G. Wang, M. Z. A. Bhuiyan, and W. Zhang, "MAPP: A modular arithmetic algorithm for privacy preserving in IoT," in *Proc. IEEE Int. Symp. Parallel Distrib. Process. Appl. IEEE Int. Conf. Ubiquitous Comput. Commun.*, 2017, pp. 897-903.
- [3] Imoize, A. L., Adedeji, O., Tandiya, N., & Shetty, S. (2021). 6G enabled smart infrastructure for sustainable society: Opportunities, challenges, and research roadmap. *Sensors*, 21(5), 1709.
- [4] Yang, B., Cao, X., Xiong, K., Yuen, C., Guan, Y. L., Leng, S., ... & Han, Z. (2021). Edge intelligence for autonomous driving in 6G wireless system: Design challenges and solutions. *IEEE Wireless Communications*, 28(2), 40-47.
- [5] Nidamanuri, J., Nibhanupudi, C., Assfalg, R., & Venkataraman, H. (2021). A progressive review: Emerging technologies for ADAS driven solutions. *IEEE Transactions on Intelligent Vehicles*, 7(2), 326-341.
- [6] Fragkos, G., Lebien, S., & Tsiropoulou, E. E. (2020). Artificial intelligent multi-access edge computing servers management. *IEEE Access*, 8, 171292-171304.
- [7] Kaur, P., Kumar, M., & Bhandari, A. (2017). A review of detection approaches for distributed denial of service attacks. *Systems Science & Control Engineering*, 5(1), 301-320.
- [8] Ameen, N., Tarhini, A., Reppel, A., & Anand, A. (2021). Customer experiences in the age of artificial intelligence. *Computers in Human Behavior*, 114, 106548.
- [9] Wu, Y. J., Hwang, P. C., Hwang, W. S., & Cheng, M. H. (2020). Artificial intelligence enabled routing in software defined networking. *Applied Sciences*, 10(18), 6564.
- [10] Li, M., Zhu, L., Zhang, Z., Lal, C., Conti, M., & Martinelli, F. (2021). Privacy for 5G-supported vehicular networks. *IEEE Open Journal of the Communications Society*, 2, 1935-1956.
- [11] Talpur, A., & Gurusamy, M. (2021). Machine learning for security in vehicular networks: A comprehensive survey. *IEEE Communications Surveys & Tutorials*, 24(1), 346-379.
- [12] Zuo, Y., Guo, J., Gao, N., Zhu, Y., Jin, S., & Li, X. (2023). A survey of blockchain and artificial intelligence for 6G wireless communications. *IEEE Communications Surveys & Tutorials*.
- [13] Priya Kohli, Sachin Sharma, Priya Matta, Secured Authentication Schemes of 6G Driven Vehicular Communication Network in Industry 5.0 Internet-of-Everything (IoE) Applications: Challenges and Opportunities, 2022 IEEE 2nd International Conference on Mobile Networks and Wireless Communications (ICMNWC).
- [14] Huang, J., Fang, D., Qian, Y., & Hu, R. Q. (2020). Recent advances and challenges in security and privacy for V2X communications. *IEEE Open Journal of Vehicular Technology*, 1, 244-266.
- [15] Xia Feng, Liangmin Wang, PAU: Privacy Assessment method with Uncertainty consideration for cloud-based vehicular networks, *Future Generation Computer Systems* 96 (2019) 368-375.
- [16] Wagan, A. A., Mughal, B. M., & Hasbullah, H. (2010, February). VANET security framework for trusted grouping using TPM hardware. In *2010 Second International Conference on Communication Software and Networks* (pp. 309-312). IEEE.
- [17] Sumra, I. A., & Hasbullah, H. B. (2015, February). Using TPM to ensure security, trust and privacy (STP) in VANET. In *2015 5th national symposium on information technology: towards new smart world (NSITNSW)* (pp. 1-6). IEEE.
- [18] Chatterjee, S., & Das, S. (2015). Ant colony optimization based enhanced dynamic source routing algorithm for mobile Ad-hoc network. *Information sciences*, 295, 67-90.

*Edited by:* Venkatesan C

*Special issue on:* Next Generation Pervasive Reconfigurable Computing for High Performance Real Time Applications

*Received:* Sep 26, 2023

*Accepted:* Dec 5, 2023



## SUSTAINABLE DEVELOPMENT IN MEDICAL APPLICATIONS USING NEURAL NETWORK ARCHITECTURE

SHUYI JIANG\*

**Abstract.** The purpose of this research is to propose a methodology utilizing machine learning techniques to support medical organizations in effectively managing risks. Specifically, the study aims to connect social media data to identify and assess potential threats, ultimately enabling healthcare management to make informed decisions for their organizations and clients. The research employs machine learning algorithms to analyze user-generated content on social media platforms, generating comprehensive visual representations of various risk categories and their magnitudes. Additionally, the study utilizes data simplification techniques, including categorization, to streamline data processing and enhance overall efficiency. A computational framework is also developed, incorporating closed-form connections for threat assessment and evaluation. The study further empirically analyses the Consumer Value Stores (CVS) established for medical care in the United States. The findings reveal that prevalent threats within the lower quartile of client messages about CVS services include operational, financial, and technological risks. The severity of these risks is distributed among high risk (21.8%), moderate risk (78%), and minimal risk (0.2%). The research also presents several metrics to demonstrate the robustness of the proposed framework, confirming its effectiveness in effectively identifying and addressing potential threats. This research provides insights that can help healthcare management make informed decisions and foster a safer and more secure environment for their organizations and the people they serve.

**Key words:** Conversational Interpretation, Artificial Intelligence, Internet-based Information Mining, Medical Analytics, Risk Administration.

**1. Introduction.** Energy usage is experiencing a notable shift, moving away from traditional fossil fuel reliance and adopting alternative energy sources. Consequently, it becomes crucial to investigate the potential correlation between decreased electricity consumption and the enhancement of Green Total Factor Productivity (GTFP). This research explores various technical methodologies that can effectively merge the dynamic relationship between energy conservation and GTFP. The interplay of Artificial Intelligence (AI) and the natural resource market profoundly impacts GTFP in a bidirectional manner. Focusing on China as a case study and employing a pertinent mathematical model, the research delves into how AI fosters environmentally sustainable economic growth. AI's impact on carbon intensity varied significantly across various industries and developmental stages, underscoring notable strides in decreasing carbon intensity during the Eleventh Five-Year Plan compared to the previous period [1].

The sectors reliant on labour and technology are predicted to witness a more significant reduction in carbon intensity attributable to AI and the dynamics of the natural resources market, in contrast to enterprises heavily dependent on capital. Furthermore, this study anticipates that the nonlinear impact of machine learning and the environmental marketplace on the overall productivity of green factors will continue to follow a "U" shaped trajectory shortly. This research formulates a quadratic equation to characterize the intricate relationship between artificial intelligence and sustainable total factor production, drawing upon the latest advancements and research in AI [2].

The impact of artificial intelligence on the overall productivity of green industries exhibits a mixed character, encompassing both advantageous and adverse effects. Given the potential complexities associated with the productivity of various factors within green industries, this study examines AI's nonlinear influence on green total factor efficiency. To achieve this, the research employs a comprehensive approach, integrating a regressive dynamic panels regression model and a differential technique of moments alongside systems GMM, considering the inherent complexities of the system dynamics [3].

---

\* College of Computer Science and Technology, Changchun University of Science and Technology, Changchun, 130012, China (shuyijiang6@163.com)

The authors investigate the potential impact of artificial intelligence on carbon dioxide emissions within China's industrial sector, commencing from 2006 and extending through 2018. Remarkably, a discernible correlation is observed between the increased deployment of manufacturing robots and a corresponding reduction in carbon intensity. However, our findings suggest that artificial intelligence, on the whole, exerts a detrimental influence on China's greenhouse gas emissions, thereby emphasizing AI's multi-layered role in sustainable industrial practices [4].

The initial empirical evidence substantiates artificial intelligence's positive impact on carbon emissions. Notably, the findings underscore the variability of AI's influence on carbon intensity, contingent upon the specific developmental stage and industry. The efficiency of AI in limiting greenhouse gas emissions in China has exhibited marked improvement, coinciding with the nation's increased capacity for technological absorption and the production of a skilled robotic workforce. Furthermore, in contrast to enterprises heavily reliant on capital, sectors emphasizing labour and technological integration are composed to witness a more pronounced reduction in their environmental footprint, owing to the integration of AI. This underscores the differential impact of AI adoption across various industrial sectors, explaining the intricate relationship between technological advancements and environmental sustainability [5].

The paper indicates a significant influence of the six variables considered, delineating their crucial roles in the potential establishment of smart towns in Nigeria. Particularly, factors about the environment, technological advancements, societal dynamics, and legal frameworks are progressively gaining prominence, underscoring their critical contributions to developing and realizing smart town initiatives in the Nigerian context. Applying a comprehensive fuzzy synthesis approach offers a realistic perspective on the challenges that necessitate resolution before the aspiration of establishing smart cities in developing nations can be actualized. This lays a robust theoretical foundation for future research endeavours focused on implementing smart cities in developing countries, particularly Africa, where environmental and socioeconomic conditions resemble those observed in the Nigerian context. It notably establishes a firm conceptual framework for further exploration in this domain [6, 7].

The nation faces many pressing challenges that demand immediate attention before any meaningful strides can be made toward realizing the potential of urban regeneration. These challenges encompass elevated rates of development, growing populations, deficient foundational infrastructure, socioeconomic inequalities, inadequate legislative frameworks, financial instability, and governance deficiencies. Therefore, the insights collected from this research hold valuable implications for governmental authorities and other stakeholders responsible for urban development, emphasizing the critical issues that necessitate resolution [8].

Policymakers must prioritize addressing these fundamental challenges to foster societal equity and promote the advancement of the nation's population. While this research has limitations, it offers valuable insights into the complex challenges of developing smart towns. The employed methodology remains critical, suggesting that future studies could benefit from an integrated approach combining the Delphi methodology for validating issues identified in the primary literature, supplemented by other statistical assessment methods. Such an approach promises a more comprehensive understanding of the intricate obstacles to advancing intelligent urban centres [9].

The smart town concept represents a pragmatic response to the multi-layered challenges modelled by global industrialization. Across the globe, technological advancements, including the Internet, the Internet of Things (IoT), artificial intelligence (AI), and data analytics, are increasingly being harnessed in communities to enhance transportation, construction, healthcare, and social services. Cities are fostering digitization and entrepreneurial growth by deploying electronic infrastructure, fostering local economies and strengthening community well-being. However, the growing smart town industry is concurrently characterized by a propagation of fragmented smart city marketplaces and initiatives, engendering concerns surrounding governance, environmental coherence, and operational consistency [10].

In pursuing sustainable development, the enhancement of human health must be intricately woven into the fabric of societal and technological progress, along with environmental preservation. Achieving this necessitates the establishment of resilient frameworks and processes and meticulous risk management to mitigate potential threats. Administrative challenges, financial risks, and technology-related perils represent only a fraction of the myriad hazards medical organizations might face, particularly those operating within the digital era. Therefore, providing exceptional services for individuals and businesses necessitates the adoption of cutting-edge

methodologies, such as machine learning, to effectively detect, assess, and address these challenges [11].

Machine learning and other AI-driven techniques are frequently utilized for textual analysis and semantic extraction. The Natural Language Processing (NLP) field emerged at the convergence of semantics and AI during the mid-20th century. Initially distinguished from Information Retrieval (IR), NLP effectively sorts and retrieves substantial text volumes through highly scalable statistical algorithms. Leveraging NLP for identifying medical conditions within digital medical records has proven instrumental in addressing multiple healthcare challenges, including monitoring diseases associated with medical care. For instance, inpatient admission rates for individuals with Chronic Obstructive Pulmonary Disease (COPD) serve as an indicator of a healthcare system's efficacy [12].

Social media's increasing influence and rapid expansion significantly impact the healthcare industry. The Internet has notably revolutionized patient interactions, providing a platform for discussing everything from support groups and the latest research to prescriptions and healthcare providers. Consequently, analyzing customer engagements on social media platforms enables medical organizations to gain comprehensive insights and improve their competitive edge. Furthermore, integrating advanced business data solutions supports optimizing risk management within healthcare facilities [13].

This research introduces an intelligent language processing framework that harnesses social media as a sophisticated data repository to enhance risk management practices within medical organizations. Moreover, the model incorporates categorization techniques and extensive data analysis methodologies to streamline data processing, facilitating more precise information handling. The study presents an algebraic overview of this framework and formulates risk analysis, identification, and assessment equations. A case study of the CVS Pharmacy Foundation in the United States is included, providing guidance on identifying and evaluating potential organizational threats. Finally, various performance indicators are scrutinized to validate the efficacy of the proposed framework [14].

**2. Related Works.** Transforming a city into a smart town is a complex and extensive undertaking, demanding the active engagement of multiple stakeholders and the ability to assess the potential of various novel computing technologies for enhancing a wide array of municipal operations. Consequently, the leadership and management of the smart city bear a significant responsibility. The study's proposed Smart City Conceptual Framework (SCCF) aims to assist municipalities in achieving this ambitious objective [15].

The Smart City Competency Model (SCCM) evaluates smart cities from administrative and technical perspectives. It equips participants with the tools to steer their cities toward a data-driven, technology-backed smart city transformation. In developing the SCCM, considerations are given to aspects encompassing strategy, internet integration, governance, and stakeholders. These core elements are supplemented by sub-components, collectively establishing robust connections and fostering a systematic approach to smart city planning, development, and implementation [16].

To more effectively manage the complex ecosystems associated with smart cities and the rapid evolution of digital technologies within urban landscapes, the study introduces and discusses the SCCM framework. This framework aims to address the resource constraints encountered by cities, which often lead to the failure of pilot smart city initiatives post-initial funding. The primary objectives of the SCCM involve facilitating the formulation of a long-term vision and plan for smart city practitioners, managing diverse interactions between stakeholders and digital infrastructures, and enabling the comprehensive assessment of associated risks and costs [17].

Central to the SCCM are the four key elements and their corresponding sub-components: strategy, technology, governance, and stakeholders. Together, the SCCM framework and its constituent components provide crucial linkages and a comprehensive approach to fostering the development of smart city initiatives and environments while removing barriers to establishing new enterprises and generating value within smart city ecosystems. Enhancing social responsibility depends upon integrating financial, social, and environmental perspectives. Within the context of smart city endeavours, a key objective frequently highlighted is enhancing residents' quality of life. Expanding the SCCM to encompass social and political dimensions makes developing environmentally sustainable and resilient smart cities attainable. Emphasizing the importance of indicators for monitoring and analyzing the multifaceted activities constituting a smart city project is crucial for ensuring its successful implementation and management [18].

In healthcare, the decision-making process is intricate, necessitating input from multiple stakeholders, and successful implementation relies on the assurance of long-term feasibility. The growing utilization of machine learning in healthcare has significantly improved medical and surgical decision-making by leveraging clinical data and statistics. However, the sustainable integration of AI-based decision systems in healthcare demands careful consideration of various factors. This study evaluates sixteen critical sustainability factors for integrating AI applications into medical decision-making, incorporating insights from 34 pertinent specialists in the Bangladeshi healthcare sector. The study ranks these factors based on expert opinions, highlighting areas of consensus and divergence [19].

The research applies data clustering techniques to categorize the variables into three distinct groups, each with essential implications for ensuring sustainability. The findings provide valuable insights for healthcare practitioners, aiding them in making informed decisions regarding integrating AI-based technologies in the healthcare sector of developing countries. The study acknowledges the potential bias introduced by respondents' subjective evaluations and employs various strategies to mitigate such bias. Future research efforts can benefit from a broader range of expert perspectives and additional metrics to ensure the investigation's objectivity [20].

In summary, this research underscores the current state of AI-driven applications in the medical sector, discussing their implications and challenges. It emphasizes the need for thoughtful approaches and strategic planning to fully connect the potential of AI, ensuring its effective integration in healthcare operations. The findings highlight the transformative impact of AI in the healthcare sector and advocate for continued exploration and implementation of AI-driven solutions to enhance overall healthcare outcomes and services.

### 3. III.Materials & Methods.

**a) Mathematical Model** This section will offer an algebraic elucidation of the proposed model, yielding several fundamental closed-form equations that can be employed for risk assessment, risk identification, and risk evaluation. The following segment will initially outline the fundamental notations in the "Abbreviation" section. Consider the postings made by individuals and others passionate about a particular medical institution as  $Q_1, Q_2, Q_3, \dots, Q_n$ . The  $n$ , in this case, represents the total amount of threads analyzed. The group encourages its patients to share their feedback on the quality of care they have received on the Internet, where it may be added to the growing body of information on a scale between  $L_1$  to  $L_n$ . We presume that every individual writes just one share, which may consist of many phrases.

$$Q = [Q_i]^T \quad (3.1)$$

The individual in need  $i$  for whom  $1 \leq i \leq n$  and  $1 \leq j \leq m$  exists has an expression value  $j$ , denoted by  $w_j$ , and  $P_i = [w_j].Q_i = [w_j]$ , where  $w_j$  is the individual  $i^{th}$  expression value.

$Q_i = [w_j]$  might store a considerable quantity of information structured in multiple manners, including nouns, verbs, and prepositions. To deal with the high complexity of the data supplied in Equation 3.1, we shall cluster the variables into distinct yet connected categories. The following equation may be used to find  $S$ :

$$S = \max[Ky] \quad (3.2)$$

where  $y$  is the dataset's phrase  $y$  and  $Ky$  is a subject integer.

As a result, Equation 3.1 may be decomposed into more manageable chunks is given by:

$$Q = [Q_u]^T \quad (3.3)$$

where  $u < i$ .

The information exists in various representations, including words, predicates, terms, modifiers, etc. To reduce the amount of information, we must consider the most basic phrases: verbs, adjectives, nouns, and adverbs. Therefore, in Equation 3.1, the vector of words connected with the understanding vector has the form:

$$Q_i = [w_d], \text{ where } d \leq j. \quad (3.4)$$

A criterion that weights the basic terms based on three instances, like in the sections that follow, to analyze

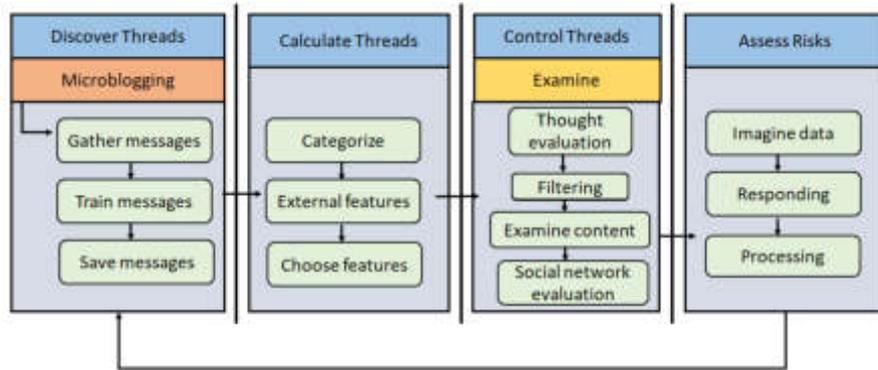


Fig. 3.1: Model of computerized risk handling in medical services

every individual’s understanding:

$$Weight = \begin{cases} 0; \text{ no risk} \\ 0.5; \text{ potential risk} \\ 1; \text{ risk} \end{cases} \tag{3.5}$$

Using the conditions presented in Equation 3.5, we get an additional closed-form expression for measuring expertise:

$$L_v = \beta \left( \frac{Verb + Adverb + Noun + Adjective}{4} \right) \tag{3.6}$$

As a result, the value of  $Q_i$  in Equation 4.4 gives rise to an additional vector whose numbers are expressed in Equation 4.6. The shape of this vector is expressed as:

$$S_r = [L_u]^T \tag{3.7}$$

The risk assessment  $Risk_{est}$  is specified as:

$$Risk_{est} = (R_1^2 + R_1^2 \sum_1^2 C_k - R_1 Pr - R_1 Pr \sum_1^2 C_k) \tag{3.8}$$

where  $Pr \leq R_1$ ,  $C_k$ =level of risk.

**b) Proposed Methodology** The following section introduces a distinctive strategy integrating the Internet, comprehensive data analysis, risk management, and healthcare. Within this framework, risks are identified, assessed, controlled, and subsequently monitored. Figure 3.1 illustrates the classification of these stages into their respective categories. Medical institutions can effectively identify potential risks by utilizing social networking platforms throughout the various phases of the risk identification process. This can be achieved by monitoring and analyzing the online discourse on platforms such as Twitter, Instagram, Facebook, and YouTube.

Engaging in online dialogues concerning disseminating specific epidemics and other health concerns is essential for promoting awareness and conducting risk evaluations. This stage encompasses three separate procedures: information gathering (utilizing APIs, crawlers, and surveys), data scrubbing (eliminating erroneous data, including misleading information, incomplete details, and conflicting data), and data retention (utilizing databases constructed with NoSQL and Apache Hive, among other extensive storage facilities). The initial step in this stage involves data collection through APIs. Subsequently, the data cleansing process follows, and finally, data storage is executed.

During the risk assessment phase, examining online user behaviours can offer insights into potential hazards for medical enterprises. The primary objective is data collection. Aside from the number of hashtags, this type of engagement can be observed through the number of shares, likes, and comments generated for specific posts. Therefore, analyzing users' mindsets and anticipating the likelihood of adverse outcomes is imperative while capturing behavioural patterns from the amassed data. This delineates how, where, to what extent, and for how long hospital patients are frequently exposed to potential risks. Data classification, feature extraction, and feature selection constitute the three sub-steps within this segment.

A wealth of audiovisual content on social media platforms, including text, images, videos, and audio files, can be abundant. These records might be organized, partially structured, or unstructured. The classification process leverages various techniques, including Quantum Support Vector Machines, a MapReduce-based k-Nearest Neighbors approach, and Composite Artificial Neural Networks to manage big data's significant scale and diversity. This involves organizing the data into relevant subsets. Progressing to the subsequent stages necessitates the initial determination of categories from the preceding step, entailing the identification of multiple attributes and aggregating the features within each category.

Feature extraction transforms input data, including words and images, into attribute-value pairs. This enhances the compatibility of features for utilization in computational learning techniques. Leveraging a set of ordered data, the N-Gram, Lexicon-based attributes, Bag-of-Words approach, and Principal Component Analysis (PCA) are employed to generate the resulting metrics, thereby reducing the overall number of variables present in the initial dataset. The objective of attribute selection is to streamline data complexity and dimensionality by reducing the number of features within the set of extracted attributes, achieved through various methods such as Chi-Square (CHI) or Information Gain (IG).

The monitoring phase incorporates various analyses, including qualitative, diagnostic, predictive, and prescriptive, enhancing decision-making capabilities by fostering a deeper understanding of pertinent factors. Descriptive analysis is the initial step in data processing, encompassing the collection of essential archival data necessary for generating actionable insights.

The segment dedicated to the results of the experiments can also facilitate the computation of the proportion of specific findings relative to the entirety. Many techniques, including data mining, association, and exploration, can be harnessed to elucidate the diagnostic analysis process, rendering it more intricate research. Through statistical analysis, data can be transformed into valuable and actionable insights. Predictive analysis is ideally conducted in stages, with pivotal components encompassing transactional targeting, decision evaluation, and predictive forecasting. By scouring the landscape for actionable insights, predictive analytics illuminates new opportunities and helps mitigate potential threats.

Stakeholders and investors associated with medical enterprises now receive valuable guidance on current risks and mitigation strategies, thanks to the illustration of insights gleaned from the preceding phase. The anticipation is that appropriate measures will be taken once feedback is received. The effectiveness of these actions will hinge on how they are executed, thus underscoring the iterative and closed-loop nature of risk management in identifying and controlling risks.

**4. Experimentation & Results.** A prominent healthcare organization for CVS, headquartered in the United States, has established itself as a global leader in the medical industry. This fact led us to determine it as an ideal candidate for a comprehensive case study. In the middle of 2022, we employed the Twint tool to extract over 28,000 retweets, focusing on the geographical distribution of individuals.

**a) Outcomes and Effectiveness Evaluations** While scrutinizing the information pool for potential threats, we conducted an initial threat assessment. Each tweet's potential threat was evaluated using a threat assessment approach. The analysis revealed that the three most frequently mentioned locations were pharmacies, healthcare facilities, and stores, prompting us to concentrate on identifying the risks associated with these areas. A comprehensive lexicon was compiled, comprising over 8000 risk-related words and phrases, such as "hacking," "malpractice," and "dangerous."

Within this lexicon, the pharmaceutical category encompassed more than 6,000 messages, the healthcare category contained over 4,300 messages, and the retail category accounted for more than 2,500 tweets for risk evaluation purposes. The findings indicated that approximately 75% of the samples exhibited no concerns, while 25% or more of the specimens indicated the presence of potential hazards. Figure 4.1 presents the monthly

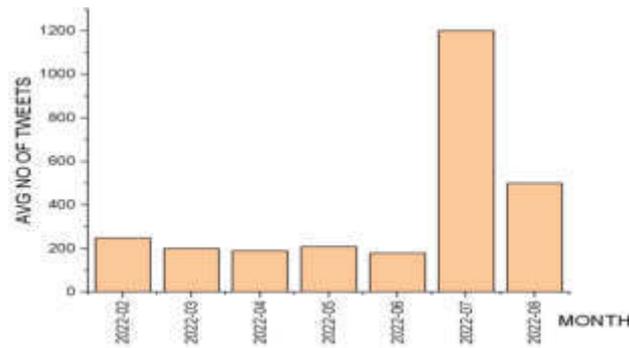


Fig. 4.1: The breakdown of tweets during six months

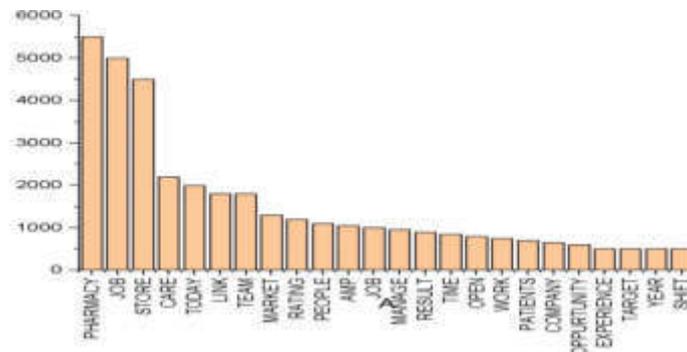


Fig. 4.2: Top-Frequency Keywords in the Information Dispersion

distribution of tweeted messages, while Table 4.2 provides several examples of comments indicating the absence of threats, the presence of potential risks, or identified risks.

The outcomes of employing Equation 3.2 to identify the most commonly occurring terms in the analyzed data are depicted in Figure 4.2. A range of performance metrics were employed to authenticate the results, including accuracy, precision, recall, and the F1 score. Given the challenges associated with annotating 28,000 tweets, we resorted to the principles of sampling statistics for assistance. To ascertain the required sample size, we applied the following formula:

$$t = \left( \frac{RT}{R - 1 + T} \right) \tag{4.1}$$

In this Equation, "T" stands for the sample proportion, and "R" stands for the error ratio. Because we have decided that the degree of conviction should be 96%, the Z-score will be the same as 1.95. In addition, the median divergence, equal to 50%, and the degree of erroneousness, equal to 5.5%, are identical. Consequently, the number of messages in the collection equals 378 when the initial values are substituted into Equation 2.9. The breakdown of the sample sizes into their respective categories is shown in Table 4.1, which can be seen below, and Figure 4.3 defines the distribution of risk categories for pharmacy as per model performance.

Now that the number of participants has been determined, we can annotate it and calculate the outcome metrics in Table 4.2. Figure 4.4 shows the model performance in risk category Vs storage.

**b) The Complicated Nature of the Suggested Approaches**

We zeroed down on the most important processes to determine the time commitment required by each suggested algorithm. The cost of the first method, which can be found in method 1, has a continuous loop. A second loop is included within this loop, and it is the one that moves across the rows of the matrix Qi. The

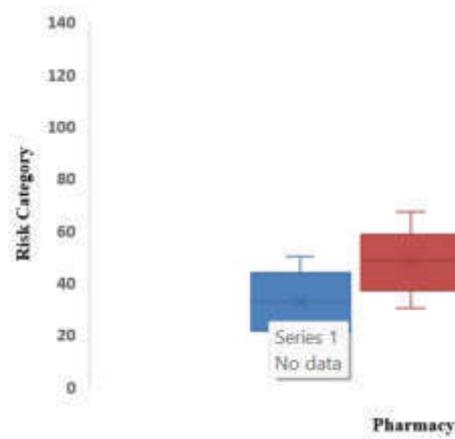


Fig. 4.3: Risk category vs. Pharmacy dataset

Table 4.1: Sample size dispersion by responsive

Data set size	Storage	Care	Pharmacy	Risk category
296	82	94	120	Zero Risk
11	3	4	4	Prospective Risk
119	22	54	43	Risk
426	107	152	147	Total

Table 4.2: Risk assessment metrics for success

S=426	Accuracy(%)	Recall(%)	F1 Score(%)	Precision(%)
Zero Risk	113	96	90	84
Prospective Risk	35	74	43	84
Risk	78	91	82	84

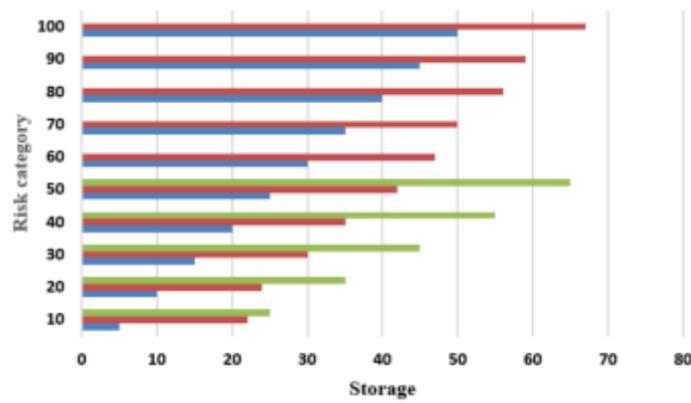


Fig. 4.4: Risk Category Vs Storage

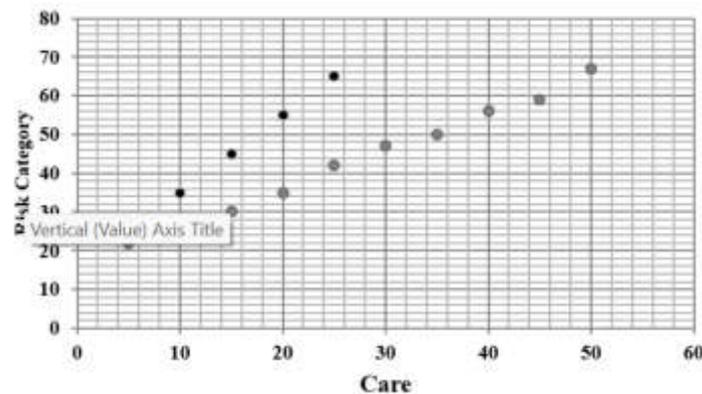


Fig. 4.5: Risk category Vs care

resulting matrix has a fixed number of rows, denoted by  $u$ , and columns, denoted by  $c$ , with the property that  $u_i \leq j$  (refer to Equation 3.2 and Equation 3.3). Therefore, the expenditure of this matrix is denoted by the notation  $O(V_c)$ .

This approach involves looking for a word using two lexicons, one for risk keywords and another for potential phrases, where the expenditures are  $O(n^2)$ , and some essential if circumstances, where their costs are  $O(1)$ . The first vocabulary is for risk phrases, while the second is for potential keywords. As a result, the total cost of the first method is denoted by the notation  $O(n^2)$ . Similarly, we utilized two vocabularies for each mechanism. Figure 6 defines the risk category Vs care. In this method, we used one vocabulary for risk terms and another for risk categories, one for risk words and the other for risk evaluations. Therefore, the time required for each of them to complete is  $O(n^2)$ .

**5. Conclusion.** The fundamental aim of this study is to employ an automatic language translation approach to detect and evaluate potential risks prevalent within medical facilities. In pursuit of this goal, the research capitalized on using social media as a contemporary and rich data source. The categorization techniques are strategically applied to streamline and manage the complexity inherent in the collected data, ensuring a more efficient and insightful analysis process. A significant outcome of the study is the derivation of a series of closed-form algebraic equations instrumental in enabling a comprehensive evaluation of the identified risks, facilitating their accurate assessment and subsequent management. A meticulous and detailed research is conducted focusing on the operations of the CVS institution, a prominent player in the American healthcare sector. This research provided a rich and informative exploration, analysis, and evaluation of the possible threats that could impact the institution's functioning and the well-being of its stakeholders. Also, calculating various essential metrics offers valuable insights into the proposed model's predictive performance. The research envisions incorporating additional methodologies. It further studies the precision and depth of the obtained findings, aiming to enhance the overall effectiveness and applicability of the proposed risk assessment model in healthcare management.

#### REFERENCES

- [1] Chang, L., Taghizadeh-Hesary, F., & Mohsin, M. (2023). Role of artificial intelligence on green economic development: Joint determinates of natural resources and green total factor productivity. *Resources Policy*, 82, 103508.
- [2] Wang, Y. (2023). Digital Economy, Technical Innovation and China's Green Total Factor Productivity Growth. *International Journal of Computational Intelligence Systems*, 16(1), 92.
- [3] Song, M., & Li, H. (2020). Total factor productivity and the factors of green industry in Shanxi Province, China. *Growth and Change*, 51(1), 488-504.
- [4] Wu, Z., Zeng, C., Huang, W., Zu, F., & Chen, S. (2022). Convergence of green total factor productivity in China's service industry. *Environmental Science and Pollution Research*, 29(52), 79272-79287.

- [5] Vergragt, P. J., & Brown, H. S. (2007). Sustainable mobility: from technological innovation to societal learning. *Journal of Cleaner Production*, 15(11-12), 1104-1115.
- [6] Okewale, R. A., & ATOBATELE, A. J. (2022). Smart Cities and Socio-Economic Development in Nigeria: Evidences from some selected Countries. *Annals of Spiru Haret University. Economic Series*, 22(2).
- [7] Mboup, G., & Oyelaran-Oyeyinka, B. (2019). Relevance of smart economy in smart cities in Africa. *Smart economy in smart African cities: Sustainable, inclusive, resilient and prosperous*, 1-49.
- [8] Aghion, P., Bacchetta, P., & Banerjee, A. (2004). Financial development and the instability of open economies. *Journal of Monetary Economics*, 51(6), 1077-1106.
- [9] Ghadami, N., Gheibi, M., Kian, Z., Faramarz, M. G., Naghedi, R., Eftekhari, M., ... & Tian, G. (2021). Implementation of solar energy in smart cities using an integration of artificial neural network, photovoltaic system and classical Delphi methods. *Sustainable Cities and Society*, 74, 103149.
- [10] Rjab, A. B., & Mellouli, S. (2021). Smart cities in the era of artificial intelligence and internet of things: promises and challenges. *Smart cities and smart governance: towards the 22nd century sustainable city*, 259-288.
- [11] Verdejo, Á., Espinilla, M., López, J. L., & Melguizo, F. J. (2022). Assessment of sustainable development objectives in Smart Labs: technology and sustainability at the service of society. *Sustainable Cities and Society*, 77, 103559.
- [12] Sheikhalishahi, S., Miotto, R., Dudley, J. T., Lavelli, A., Rinaldi, F., & Osmani, V. (2019). Natural language processing of clinical notes on chronic diseases: systematic review. *JMIR medical informatics*, 7(2), e12239.
- [13] Wang, Y., Kung, L., & Byrd, T. A. (2018). Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations. *Technological forecasting and social change*, 126, 3-13.
- [14] Scott, D. M. (2016). United States health care system: a pharmacy perspective. *The Canadian Journal of Hospital Pharmacy*, 69(4), 306.
- [15] Khan, H. H., Malik, M. N., Zafar, R., Goni, F. A., Chofreh, A. G., Klemeš, J. J., & Alotaibi, Y. (2020). Challenges for sustainable smart city development: A conceptual framework. *Sustainable Development*, 28(5), 1507-1518.
- [16] Nikolov, R., Shoikova, E., Krumova, M., Kovatcheva, E., Dimitrov, V., & Shikalanov, A. (2016). Learning in a smart city environment. *Journal of Communication and Computer*, 13(7), 338-50.
- [17] Salerno, S., Nunziante, A., & Santoro, G. (2014). Competences and knowledge: Key-factors in the smart city of the future. *Knowledge Management & E-Learning*, 6(4), 356.
- [18] Angadi, R. (2019). Introduction to Smart City and Agricultural Revolution: Big Data and Internet of Things (IoT). In *Smart Cities and Smart Spaces: Concepts, Methodologies, Tools, and Applications* (pp. 135-176). IGI Global.
- [19] Tutun, S., Johnson, M. E., Ahmed, A., Albizri, A., Irgil, S., Yesilkaya, I., ... & Harfouche, A. (2023). An AI-based decision support system for predicting mental health disorders. *Information Systems Frontiers*, 25(3), 1261-1276.
- [20] Nasseef, O. A., Baabdullah, A. M., Alalwan, A. A., Lal, B., & Dwivedi, Y. K. (2022). Artificial intelligence-based public healthcare systems: G2G knowledge-based exchange to enhance the decision-making process. *Government Information Quarterly*, 39(4), 101618.

*Edited by:* Venkatesan C

*Special issue on:* Next Generation Pervasive Reconfigurable Computing for High Performance Real Time Applications

*Received:* Sep 27, 2023

*Accepted:* Nov 21, 2023



??

## GROUP INTELLIGENT CITY MOBILE COMMUNICATION NETWORK'S CONTROL STRATEGY BASED ON CELLULAR INTERNET OF THINGS

JIAZHENG WEI\*

**Abstract.** Mobile communication network optimization heavily depends on power control technology, which impacts the effectiveness of the network. This paper aims to enhance control over nonlinear mobile communication networks and achieve superior performance by applying the particle swarm optimization (PSO) algorithm in the control domain. Addressing limitations in the basic PSO algorithm, improvements are made and applied to urban mobile communication networks. The methodology involves modifying the PSO algorithm to address identified issues and applying the enhanced algorithm to communication network scenarios. Simulation results indicate that with an initial particle count of 10 and 100 iterations, the optimized values for and are 0.691 and 0.486, respectively, resulting in an objective function value of 55.514. This achievement validates the successful implementation of the optimization process for mobile communication network control. The findings reveal that the proposed grad particle swarm optimization (Grad-PSO) algorithm exhibits mobile network optimization by robust search capability and rapid convergence.

**Key words:** Grad-PSO, Particle movement, Mobile communication network, Optimal control, Internet of Things (IoT), Learning factors.

**1. Introduction.** Mobile communication technology is undergoing profound transformations due to the swift advancement of intelligent appliances and the Internet of Things. Primarily, mobile communication is controlled to sustain a rapid developmental pace regarding user count and overall service. According to the International Telecommunication Union (ITU), the tally of mobile subscribers reached nearly 7 billion by the close of 2014, with mobile broadband's growth rate consistently in the double digits. The surging demand for mobile communication network technology presents significant opportunities and challenges, fueling a surge in research and development across academic and industrial sectors, focusing on novel services, technologies, standards, and products [1].

Within this dynamic context, 5G has emerged as a prominent focal point of communication technology's evolution, demonstrating extensive application prospects. Mobile communication network users are surging, business domains are expanding, and network equipment is diversifying extensively. It compels mobile communication networks to enhance the provision of diverse services to an expanding user base, all while upholding communication quality. Consequently, the prerequisites for mobile communication-related technologies are escalating distinctly. The pivotal technology within mobile communication networks efficiently mitigates the power control technology by direct interference among users operating on adjacent or near channels due to the "far and near effect." This substantial enhancement strengthens mobile communication networks' capacity and quality [2].

The Internet of Things (IoT) is a pivotal driving force for the evolution of mobile communications. Mobile Internet has revolutionized conventional mobile communication services, furnishing users with experiences such as ultra-high definition (3D) video, augmented and virtual reality, mobile cloud, and other immersive, cutting-edge ventures. It has spurred a comprehensive metamorphosis in the information interaction mode, speeding up the rapid maturation of mobile communication technology and the industry [3].

Moreover, the IoT has expansively broadened mobile communication applications. This expansion transcends interpersonal communication, spanning into intelligent interconnections between various objects, facilitating the application of mobile communication technology across various industries and sectors. The spread of IoT

---

\*The phased achievement of the Basic Ability Improvement Project of Young and middle-aged Teachers in Guangxi in 2019, the research and application of intelligent AC charging pile based on Internet of Things code scanning, Project No. :2019KY1346. Guangxi Transportation Vocational and Technical College, Traffic Information Engineering Institute, Nanning Guangxi, 530023, China ([weijiazheng9@163.com](mailto:weijiazheng9@163.com))

applications, ranging from mobile healthcare and smart homes to industrial control, vehicular networking, and environmental monitoring, will manifest in greater abundance [4].

The profound vision of an interconnected world, termed "the Internet of Everything," is controlled to materialize as an extensive array of IoT devices interface with the network for diverse purposes. This surge in IoT-driven applications has given rise to many emerging industries, thus catalyzing the robust advancement of mobile communication technology and the industry. Against this "Internet of Everything" demand, the pressing need for massive device connectivity, diverse services, and distinct user experiences has emerged as new focal points for mobile communication research. Regarding the challenge of power control within mobile communication networks, scholars worldwide have undertaken extensive research, applying an extent of optimization algorithms to address this issue. While algorithms like genetic, ant colony, and particle swarm optimization methods have made strides in resolving power control problems, their shortcomings, namely sluggish convergence, modest precision, and susceptibility to local optimization, adversely impact real-time performance and accuracy. In response to these challenges and aligning with the IoT's stipulations for high reliability, low latency, and reduced energy consumption, incorporating cache resources within the IoT is paramount in efficiently managing the escalating IoT traffic [5, 6].

The paper's section organization follows a structured progression to investigate the control method for group intelligent city mobile communication networks using the Cellular Internet of Things (CIoT). The literature review surveys relevant aspects of mobile communication networks, CIoT, and swarm intelligence algorithms are discussed in Section 2. The proposed methodology outlines the proposed control approach, integrating PSO within CIoT, explained in Section 3. The simulation and results are demonstrated in Section 4, which details the method's effectiveness through experimental comparisons and performance analysis. The conclusion summarizes findings, contributions, and implications, emphasizing the significance of the research in shaping the future of group intelligent city networks based on cellular IoT.

**2. Literature review.** Many experts and scholars have exhibited a pronounced interest in swarm intelligence, delving into novel approaches to tackle conventional intricate quandaries. Their pursuit involves scrutinizing social insects' societal conduct, precisely, the collaborative food-finding endeavours of simple insects. Rooted in swarm intelligence, these diligent minds have introduced an array of algorithms, with the ant colony algorithm and PSO emerging as the most emblematic exemplars [7].

The PSO algorithm was introduced by Kennedy and Eberhart in 2000, drawing inspiration from the seeking behaviours of birds. It has evolved into a strong tool for nonlinear continuous optimization, combinatorial optimization, mixed-integer nonlinear optimization, and various other optimization challenges due to its straightforward process, limited parameter requirements, simple algorithmic structure, and ease of implementation. Nevertheless, nonlinear PSO possesses shortcomings such as insensitivity to environmental changes and susceptibility to local minima. In recent years, scholars have devised enhanced algorithms grounded in nonlinear PSO. Key advancements encompass parameter tuning, particle diversity selection, population structure determination, and amalgamation with other intelligent techniques. Relative to counterparts like ant colonies and genetic algorithms, PSO's advantage lies in its fewer adjustable parameters; however, their meticulous selection significantly influences accuracy and efficiency [8].

Three universally applicable principles are introduced for selecting population size, iteration count, and particle velocity to enhance nonlinear PSO's effectiveness. Additionally, experts have incorporated PSO with other methods to overcome the local optima problem through population diversity control, balancing particle attraction and repulsion processes. Further innovations include an adaptive PSO to navigate spatial changes in dynamic systems, along with Qiu's modulation strategy for mobile communication systems [9].

The theoretical foundation for PSO enhancements and applications remains underdeveloped. Particle swarm optimization parameters are confined mainly to experimental realms, lacking comprehensive and well-defined understanding. Therefore, exploring nonlinear particle swarm algorithms holds profound importance for comprehending their internal mechanics and expanding their scope. Cellular networks have evolved as a foundation in mobile communications, offering extensive coverage and reliable communication. Projections by Qualcomm indicate that global IoT connectivity will surpass 5 billion by 2025, around diverse applications from wearable devices to environmental monitoring. The propagation of intelligent devices connecting to cellular networks positions them as the primary infrastructure for the Internet of Things [10].

The enhancement and practical utilization of the PSO algorithm suffer from a shortage of theoretical robustness. Parameters within the PSO algorithm remain within experimental realms, lacking substantial and well-defined conceptualization. Thus, as an emerging intelligence paradigm, the nonlinear particle swarm algorithm is essential for investigating its intrinsic mechanisms and expanding its application spheres. This IoT growth extends across smart cities, transportation, environmental monitoring, and medical care, encompassing an array of facets, from intelligent wearables and water/electricity meters to smart infrastructure like utility hole covers and vehicular terminals. As such, many intelligent endpoints will interface with the network, thus establishing cellular networks as the primary backbone for the Internet of Things [11].

The authors focus on crafting and executing machine learning methodologies to enhance smart cities' data processing efficiency, decision-making, and resource management capabilities. The research is anticipated to explore a range of machine learning algorithms, including neural networks, support vector machines, clustering, and deep learning models. These techniques can be analyzed for their potential utility across intelligent city domains such as traffic management, energy optimization, waste management, public safety, and healthcare [12].

This study conducts an exhaustive survey on the potential applications of utilizing 5G network-based IoT for demand response within smart grids. The investigation scrutinizes how this innovative strategy can augment grid efficiency and responsiveness, contributing to a more sustainable and adaptable energy ecosystem. The research gap becomes evident in the need for an all-encompassing, interdisciplinary approach that bridges the theoretical prospects of 5G network-based IoT for demand response in smart grids with tangible considerations, economic feasibility, regulatory hurdles, and human-centred aspects. Such an approach would yield a more comprehensive grasp of the authentic potential, limitations, and prerequisites essential for effectively merging these technologies, ultimately shaping the future energy management trajectory [13].

An adequate examination of the pragmatic challenges and constraints of deploying such a system within intricate and dynamic urban settings is lacking. Despite proposing an inventive IoT-based method for accident detection and reporting in smart cities, the study frequently neglects potential hindrances associated with real-world implementation. Fundamentally, although the envisioned IoT-powered accident detection and reporting system exhibits potential, additional research is imperative to bridge the disparity between theoretical concepts and pragmatic complexities. This endeavour is essential to ensure the system's efficacy, dependability, and smooth adjustment within smart city environments [14].

The research fails to thoroughly examine the challenges and constraints of merging big data analysis and deep learning techniques for constructing digital twins in smart cities. Although the suggested approach exhibits potential for supporting smart city planning via digital twins, additional research is imperative to tackle the noted deficiencies. This encompasses comprehending hurdles in data integration, ensuring model adaptability, considering resource ramifications, fostering interdisciplinary collaboration, and addressing ethical considerations. Addressing these gaps will play a pivotal role in harnessing the full potential of big data analysis and deep learning for the inception and application of digital twins in smart urban environments [15].

**3. 3. Proposed Grad-Particle Swarm Optimization algorithm.** The PSO algorithm characterizes each solution within an optimization problem as a "particle." The fitness values of these particles derive from the objective function under optimization. Furthermore, individual particles possess velocities, prompting them to trail the presently optimal particle across the solution space during the iterative search until the ultimate solution surfaces.

Figure 3.1 illustrates the optimal control flowchart for a group-intelligent mobile communication network founded upon the cellular IoT paradigm. Each particle refines its position by monitoring two "extremes" throughout each iteration. The particle's self-derived optimal solution is termed the individual extremum.  $p_{best}$  While the prevailing optimal solution for the entire population is known as the global extremum  $g_{best}$  Particles must continually update both their velocity and position, a process governed by Equations 3.1 and 3.2.

$$v_i = \omega v_{i-1} + c_1 \times r_1 \times [p_{best} - x_{i-1}] + c_2 \times r_2 \times [g_{best} - x_{i-1}] \quad (3.1)$$

$$x_i = x_{i-1} + v_i \quad (3.2)$$

In the context of this representation,  $v_i$  and  $v_{i-1}$  stand for the current and preceding particle movement speeds, correspondingly. Likewise,  $x_i$  and  $x_{i-1}$  denote the current and former particle positions, respectively. The

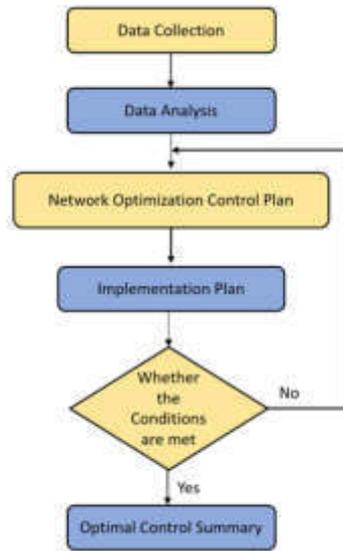


Fig. 3.1: Flowchart for a group-intelligent mobile communication network

symbol  $\omega$  signifies inertial weights, while  $p_{best}$  and  $g_{best}$  indicate individual and global extrema, respectively. The parameters  $c_1$  and  $c_2$  encapsulate learning factors, typically adopting equal values like  $c_1 = c_2 = 2$ . Furthermore,  $r_1, r_2$  are random numbers spanning the range from 0 to 1.

By amalgamating the operational swiftness and precision benefits inherent in traditional value optimization techniques, a novel approach seeks to enhance the convergence velocity of the PSO algorithm. To this end, the gradient method is infused into the PSO framework, culminating in a Grad-PSO algorithm fortified by a gradient search factor. This algorithm posits the existence of a global minimum for the optimization function within the domain space. It envisions a circular region "A," centred at the global minimum "g," and circumscribed by a radius "r," representing the optimal zone.

During particle movement, when far from the global optimal value, the original position updating strategy of the PSO algorithm is retained. In contrast, the gradient method governs position updates when proximity to the global optimum is achieved. This strategic integration alleviates the computational burden introduced by the stochastic particle position updates within the PSO algorithm. Upon entering the optimal region, the gradient method guides particles to converge towards the optimal position swiftly, amplifying the optimization pace.

While introducing the gradient method compromises some of the PSO algorithm's randomness and adaptability, it ensures particles within the optimal region remain confined, heightening single optimization efficiency. Consequently, the overall optimization efficiency of the algorithm is markedly enhanced. The specific velocity and position updating mechanism of the Grad-PSO algorithm is articulated through Equations 3.3 to 3.4.

$$x = x + v, f(x) > f(g) + r \tag{3.3}$$

$$x = T(x), f(x) \leq f(g) + r \tag{3.4}$$

To verify the advantages of the proposed Grad-PSO algorithm, the following mathematical problems are analyzed. The target function is  $\min f(x) = x^2 + 2x + 6$ . The constraint is  $-10 \leq x \leq 10$ . The problem is optimizing one variable function with boundary constraints [16]. The global optimal solution is  $f(x) = 5.0000$ . Matlab is used for programming, and the simulation results are shown in Figure 3.2.

The simulation experiment yields evidence that the Grad-PSO algorithm exhibits enhanced regularity within the optimization function, owing to the incorporation of the gradient search factor. This characteristic

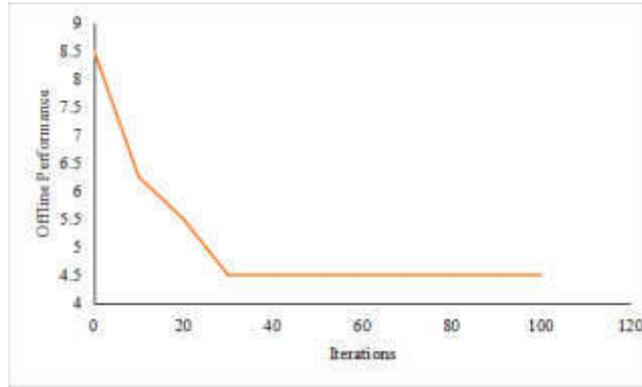


Fig. 3.2: Simulation result of optimization function of Grad-PSO algorithm

demonstrates heightened optimization efficiency and remarkable precision. In summation, it can be deduced that the Grad-PSO algorithm is a notably superior approach to function optimization. Within CDMA technology-based cellular mobile communication systems, user terminals employ a shared spectrum for uplink and downlink data transmission, inevitably leading to user interference. A prominent illustration is in broadband CDMA cellular mobile communication systems, characterized by the "near and far effect." This effect is intimately linked with channel power during user communication. Thus, effective management of the signal power of user terminals is essential to mitigate this phenomenon.

Furthermore, optimizing the transmission power of base stations plays a pivotal role. This optimization ensures that each user terminal receives an appropriate radiated power from the base station, contributing to improved overall system performance. In a mobile communication system, the capacity and efficiency of frequency spectrum utilization are directly contingent upon each user's signal power and transmission rate. A comprehensive mathematical model is devised to address power control challenges within mobile communication networks, accounting for the interplay between power control and rate control strategies.

Consider a multi-cell DS-CDMA cellular mobile communication system comprising  $N$  users, collectively sharing a spread bandwidth denoted as  $W$ . Notably, each user imposes distinct requisites concerning transmission rate, delay, and bit error rate. For analytical convenience, let's define the maximum allowable transmission power per user as  $P_{maxi}$  alongside the minimum required transmission rate denoted as  $R_{mini}$ . In addition, a designated target bit energy-to-noise ratio ( $E_b/N_0$ ) is selected to align with specific bit error rate criteria. This enables incorporating a variable transmission rate within a predetermined range to cater to individual user-imposed delay constraints.

Integral to this model is the representation of critical factors. The parameter  $h_{ii}$  captures the channel gain from user "i" to the receiver at its base station. Similarly,  $h_{ij}$  signifies the channel gain originating from user "j" to the receiving base station of user "i." The signal transmission power of user "i" is denoted as  $P_i$  or  $\gamma_i$  aligned with the target  $E_b/N_0$  requirement. Additionally, the background noise encountered at the base station receiver assumes an additive white Gaussian nature, characterized by a unilateral power spectral density of  $\eta_0$ . Central to this framework is the normalized signal-to-noise ratio at the base station, conveyed as  $E_b/N_0$ . The illustration of this scenario is encapsulated within Equation 3.5, showcasing the reception of the user signal by the base station.

$$\frac{E_b}{N_0} = \frac{W}{R_i} \cdot \frac{h_{ii} \times P_i}{\sum_{i \neq j} h_{ij} \times P_j + \eta_0 \times W} \quad (3.5)$$

$$i = 1, 2, 3, \dots, N \quad (3.6)$$

To moderate interference among users in different cells, optimization is pursued through the minimization of total transmitted power. Addressing scenarios where the system is congested and quality of service requirements

must be upheld, a priority control strategy is introduced as follows: With a commitment to maintaining high-priority services, the strategy endeavours to augment the transmission quality of low-priority services. This augmentation is achieved by carefully elevating transmission power. To organize user priorities, a coefficient  $A_i$  is introduced [17]. The optimization problem's objective function is formulated in Equation 3.6, while the constraints are articulated through Equations 3.7 to 3.9.

$$\min \sum A_i P_i \tag{3.7}$$

$$\frac{E_b}{N_0} \geq \gamma_i \tag{3.8}$$

$$0 \leq P_i \leq P_{maxi} \tag{3.9}$$

$$R_i \geq R_{mini} \tag{3.10}$$

The evaluation of the objective function's values throughout various iterations evaluates the algorithm's advancement and its rate of convergence. This iterative process of enhancement, guided by the objective function, amplifies the algorithm's effectiveness in addressing complex problems.

**4. Results and Discussion.** To simplify the calculation, the mathematical model is simplified, starting from a relatively simple case  $N=2$ . Set up  $N=2$ ,  $i=1,2$ ,  $P_1 = x_1, P_2 = x_2, R_1 = y$ . According to the actual situation of the mobile communication network, set  $W = 100MHz$ ,  $\eta_0 = 2 \times 10^{-8}$ ,  $h_{11} = 2, h_{12} = 3, h_{21} = 1.5, h_{22} = 2.5, \gamma_i = 0.8, P_{max1} = P_{max2} = 1W, R_{mini} = 50Kb/s$  weight coefficient  $A_1 = 30, A_2 = 100$ . The grad-PSO algorithm is used to optimize the power control. Its objective function is shown in Equation 4.1, and its constraint conditions are shown in Equations 4.2-4.3.

$$f(x_1, x_2) = \min(30x_1 + 100x_2) \tag{4.1}$$

$$\frac{200x_1}{y(3x_2 + 2)} \geq 0.8 \tag{4.2}$$

$$\frac{250x_1}{y(1.5x_2 + 2)} \geq 0.8 \tag{4.3}$$

where  $0 \leq x_1 \leq 1; 0 \leq x_2 \leq 1; y \geq 50$ .

The specific processing flow of the Grad-PSO algorithm is as follows.

Step 1: Set each parameter of the algorithm, such as the number of particles contained in the population, that is the size of the population . Coefficient of inertia weight  $\omega = 0.9$  and acceleration constants  $c_1 = c_2 = 2$ .

Step 2: Perform arbitrary initialization on the particles of the population (the population size is  $N$ ), and calculate and determine the fitness of all particles;

Step 3: Evaluate the fitness of each particle calculated in step 2;

Step 4: Match the fitness value of each particle with the historical best position the comparison was performed. If the current particle fitness value is better, then is updated to the current fitness value;

Step 5: For each particle, the fitness value of fitness is combined with the historical optimal position of the population  $g_{best}$ . If the best fitness value in the current population is better than the historical best  $g_{best}$ . Then update it to  $g_{best}$ .

Step 6: Update the position and velocity of each particle according to the formula;

Step 7: Calculate the performance index to see whether it meets the optimization end condition. If the condition is met, the current result is the optimal solution, and the algorithm ends. Otherwise, return to Step 3 and continue the next loop.

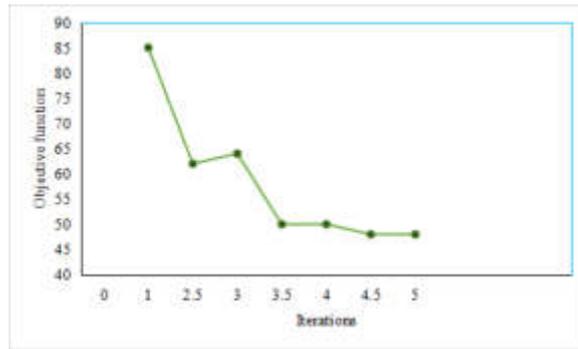


Fig. 4.1: Objective function when the initial particle number is 10

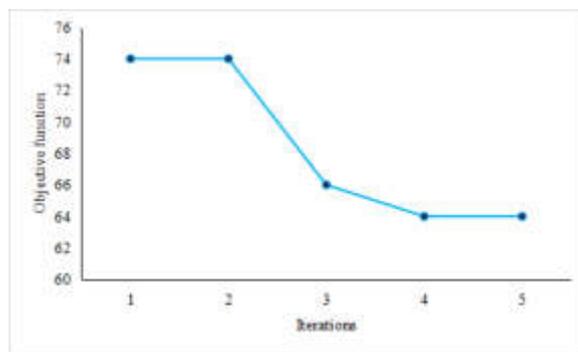


Fig. 4.2: Objective function when the initial particle number is 20

By applying the grad-PSO algorithm to solve the objective function, the optimal solution can be obtained as  $x_1 = 0.6916, x_2 = 0.4860$ . The target function is zero for  $f = 55.5140$ . Figures 4.1 and 4.2 show the changes and comparison of the offline performance curve of the Grad-PSO algorithm when the number of iterations is five, and the initial particle number is 10 and 20.

Upon comparing Figure 4.1 and Figure 4.2, it becomes evident that the Grad-PSO algorithm exhibits disparate convergence speeds and paths while undergoing the same number of iterations but with varying initial particle quantities. However, it's noteworthy that, ultimately, both scenarios attain an identical optimal solution.

Table 4.1, presented below, illustrates the shifts in particle optimal positions ( $x_1, x_2$ ) and objective function value ( $f$ ) with changes in iteration count for initial particle numbers 10 and 20, respectively. The data in Table 4.1 demonstrates that as the iteration count escalates, the Grad-PSO algorithm progressively approaches the optimal value, remaining resilient against divergence from the optimal solution due to inherent randomness. Despite their distinct initial particle counts, the convergence trajectory and pace vary under equivalent iteration counts. Nonetheless, both scenarios efficiently converge towards the optimal power control objective, swiftly realizing the optimization of the power control function.

**5. Conclusion.** This paper explored the power control principles that lead to a simplified model, viewed through the lens of joint power and rate control. The Grad-PSO algorithm, a typical group intelligence technique, finds wide application in engineering optimization challenges. Implementing the Grad-PSO in mobile communication networks enhances blind channel equalization and communication quality. The Grad-PSO algorithm claims discontinuity and differentiation, powerful search capabilities, and rapid convergence rates. As particle count rises, so does the likelihood of achieving optimal solutions. Simulations demonstrate that with an initial particle count of 10 and 100 iterations,  $x_1$  attains 0.691,  $x_2$  reaches 0.486, and the objective function

Table 4.1: Particle optimal position and objective function values with the number of iterations

Initial number of particles	10				20			
	5	20	40	100	5	20	40	100
$x_1$	0.782	0.693	0.691	0.691	0.811	0.691	0.691	0.691
$x_2$	0.416	0.488	0.486	0.486	0.547	0.486	0.486	0.486
f	59.441	55.757	55.517	55.514	62.892	55.561	55.515	55.514

value stands at 55.514, ultimately securing an optimal solution and enabling effective mobile communication network control. Furthermore, as the PSO algorithm refines its iterative processes, the probability of rapid optimal solution attainment surges.

## REFERENCES

- [1] Zhao, Y., Yu, G., & Xu, H. (2019). 6G mobile communication network: vision, challenges and key technologies. arXiv preprint arXiv:1905.04983.
- [2] Ni, Y., Liang, J., Shi, X., & Ban, D. (2019, January). Research on key technology in 5G mobile communication network. In 2019 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS) (pp. 199-201). IEEE.
- [3] Jeong, Y. S., & Park, J. H. (2019). IoT and smart city technology: challenges, opportunities, and solutions. *Journal of Information Processing Systems*, 15(2), 233-238.
- [4] Elazhary, H. (2019). Internet of Things (IoT), mobile cloud, cloudlet, mobile IoT, IoT cloud, fog, mobile edge, and edge emerging computing paradigms: Disambiguation and research directions. *Journal of network and computer applications*, 128, 105-140.
- [5] Zhang, H., Shlezinger, N., Guidi, F., Dardari, D., Imani, M. F., & Eldar, Y. C. (2022). Near-field wireless power transfer for 6G internet of everything mobile networks: Opportunities and challenges. *IEEE Communications Magazine*, 60(3), 12-18.
- [6] Liu, Y., Dai, H. N., Wang, Q., Shukla, M. K., & Imran, M. (2020). Unmanned aerial vehicle for internet of everything: Opportunities and challenges. *Computer communications*, 155, 66-83.
- [7] Gupta, A., & Srivastava, S. (2020). Comparative analysis of ant colony and particle swarm optimization algorithms for distance optimization. *Procedia Computer Science*, 173, 245-253.
- [8] Yu, M. (2019). A solution of TSP based on the ant colony algorithm improved by particle swarm optimization. *Discrete and Continuous Dynamical Systems-S*, 12(4&5), 979-987.
- [9] Zhang, X., Jia, M., Chen, L., Ma, J., & Qiu, J. (2015, December). Filtered-OFDM-enabler for flexible waveform in the 5th generation cellular networks. In 2015 IEEE global communications conference (GLOBECOM) (pp. 1-6). IEEE.
- [10] Bello, O., & Zeadally, S. (2014). Intelligent device-to-device communication in the internet of things. *IEEE Systems Journal*, 10(3), 1172-1182.
- [11] Pasolini, G., Buratti, C., Feltrin, L., Zabini, F., De Castro, C., Verdone, R., & Andrisano, O. (2018). Smart city pilot projects using LoRa and IEEE802. 15.4 technologies. *Sensors*, 18(4), 1118.
- [12] Din, I. U., Guizani, M., Rodrigues, J. J., Hassan, S., & Korotaev, V. V. (2019). Machine learning in the Internet of Things: Designed techniques for smart cities. *Future Generation Computer Systems*, 100, 826-843.
- [13] Hui, H., Ding, Y., Shi, Q., Li, F., Song, Y., & Yan, J. (2020). 5G network-based Internet of Things for demand response in smart grid: A survey on application potential. *Applied Energy*, 257, 113972.
- [14] Bhatti, F., Shah, M. A., Maple, C., & Islam, S. U. (2019). A novel internet of things-enabled accident detection and reporting system for smart city environments. *sensors*, 19(9), 2071.
- [15] Li, X., Liu, H., Wang, W., Zheng, Y., Lv, H., & Lv, Z. (2022). Big data analysis of the internet of things in the digital twins of smart city based on deep learning. *Future Generation Computer Systems*, 128, 167-177.
- [16] Yazdanpanah, F., Afsharmazayejani, R., Alaei, M., Rezaei, A., & Daneshlab, M. (2019). An energy-efficient partition-based xyz-planar routing algorithm for a wireless network-on-chip. *Journal of supercomputing*, 75(2), 837-861.
- [17] Sun, B., Ding, J., Lin, S., Wang, W., Chen, Q., & Zhong, Z. (2019). Comparison analysis on feasible solutions for LTE based next-generation railway mobile communication system. *ZTE Communications*, 17(01), 60-66.

*Edited by:* Venkatesan C

*Special issue on:* Next Generation Pervasive Reconfigurable Computing for High Performance Real Time Applications

*Received:* Sep 27, 2023

*Accepted:* Nov 21, 2023



## METHODOLOGY FOR DEVELOPING AN IOT-BASED PARKING SPACE COUNTER SYSTEM USING XNO

SUJANAVAN TIRUVAYIPATI\*, RAMADEVI YELLASIRI†, VIKRAM NARAYANDAS‡, ARCHANA MARUTHAVANAN§  
AND ANUPAMA MEDURI¶

**Abstract.** The Blockchain-IoT integration is treated as the future technological value by many adopters despite the cost or complexity involved. But, there are technological advancements brought in by communities that make solutions affordable and simple to the fact that they can be used in applications such as parking space counters. This research portrays use of XNO which is a digital currency in an alternative way to keep track of available parking spaces via IoT nodes installed at entry and exit points of parking lots. The available parking spaces data using this approach can be displayed on LED boards at the entry point of the parking lots and on a website for remote status view. An add-on for this research is the issue of entry tickets with timestamp and unique ID by using the block data during asset transfer. This research can be further enhanced for collection of parking fees by the help of IoT nodes at the exit points of the parking lots.

**Key words:** IoT, Block-chain, Block-lattice, Energy-efficient, Sensors, Cryptocurrency, Parking Space Counter.

**AMS subject classifications.** 11U99, 11Y55, 11Z05, 28E99, 28A99

**1. Introduction.** The origin of IoT dates back decades in comparison with Blockchain being coined. Long story short Tab.1.1 shows the historical technological evolution that brought in IoT and Blockchain.

Bitcoin was the first application of blockchain but unfortunately as Bitcoin matured several issues in the protocol made Bitcoin prohibitive for many applications which include: Poor scalability, High latency, Power inefficient. In 2012 an alternative consensus protocol, Proof of Stake (PoS) [12], was first introduced by Peercoin. PoS does away with the wasteful computation power competition, only requiring light-weight software running on low power hardware. Further enhancements were made when the first Directed Acyclic Graph (DAG) based cryptocurrencies [13] [14] [15] broke the blockchain mold improving system performance and security using a consensus system that provides quicker, more deterministic transactions while still maintaining a strong decentralized system. Meanwhile IoT had its own enhancements in 2014, Constrained Application Protocol (CoAP) [IETF RFC 7252] for low-power, lossy networks over UDP was introduced [16].

In order to understand the value of the modern blockchain-based solutions one should refer to the traditional method of data collection from IoT nodes by a centralized database server as shown in Fig.1.1 which also handles multiple user requests which could lead to SPOF(Single Point of Failure).

In the world of IoT applications parking systems also have a considerable market share due to growing population, vehicles and traffic. George Mason University in Fairfax, VA, faced a parking deficit due to growth forecasts and to address this, a system was designed [17] to identify commuter populations, predict ridership, and analyze utility of shuttle routes or garages as the goal is to reduce traffic congestion and emissions, ultimately

---

\*Department of Computer Science and Engineering, Maturi Venkata Subba Rao Engineering College, Osmania University, Hyderabad, Telanagana, India ([sujanavan\\_cse@mvsrec.edu.in](mailto:sujanavan_cse@mvsrec.edu.in)).

†Department of Computer Science and Engineering, Chaitanya Bharathi Institute of Technology, Osmania University, Hyderabad, Telanagana, India ([yramadevi\\_cse@cbit.ac.in](mailto:yramadevi_cse@cbit.ac.in)).

‡Department of Computer Science and Engineering, Maturi Venkata Subba Rao Engineering College, Osmania University, Hyderabad, Telanagana, India Research Scholar, Department of Information Technology, Faculty of Engineering and Technology, Annamalai University, Chidambaram, Tamil Nadu, India ([vikramnarayandas16@gmail.com](mailto:vikramnarayandas16@gmail.com)).

§Assistant Professor, Department of Information Technology, Faculty of Engineering and Technology, Annamalai University, Chidambaram, Tamil Nadu, India ([archana.aucse@gmail.com](mailto:archana.aucse@gmail.com)).

¶Department of Computer Science and Engineering, Maturi Venkata Subba Rao Engineering College, Osmania University, Hyderabad, Telanagana, India ([anupama\\_cse@mvsrec.edu.in](mailto:anupama_cse@mvsrec.edu.in))

Nano(XNO) is a cryptocurrency launched in October 2015 by Colin LeMahieu to address the Bitcoin scalability problem.

Table 1.1: A history of events that lead to the IoT-Blockchain integration.

Year	Event
1954	Birth of the transistor-based electronic computer [1]
1969	ARPANET(Advanced Research Projects Agency Network) [2] was the first wide-area public packet-switched computer network to implement the TCP/IP protocol suite developed by DARPA (Defense Advanced Research Projects Agency) for academic and research purposes. ARPANET was decommissioned in 1989 and paved the way for modern internet.
1982	A beverage machine [3] connected via the internet at Carnegie Mellon University’s School of Computer Science in Pittsburgh to remotely check if the machine was loaded with cold coke. The first IoT sensor data broadcast.
1982	Cryptographer David Chaum first proposed a blockchain-like protocol [4].
1990	A Sunbeam Deluxe Automatic Radiant Control Toaster was connected to the Internet [5], becoming the hit of the 1990 Interop. The first remote controlled actuator in IoT.
1991	Work on a cryptographically secured chain of blocks was described by Stuart Haber and W. Scott Stornetta [6].
1993	In 1992, Haber, Stornetta, and Dave Bayer incorporated Merkle trees into the design [7], which improved its efficiency by allowing several document certificates to be collected into one block.
1999	A British technology pioneer Kevin Ashton, co-founder of the Auto-ID Laboratory at MIT coined the term “The Internet of Things” [8] to describe a system where the Internet is connected to the physical world via ubiquitous sensors including RFID (Radio-Frequency IDentification).
1999	MQTT (Message Queue Telemetry Transport) a lightweight protocol over TCP suitable for IoT. MQTT is an OASIS standard and an ISO recommendation (ISO/IEC 20922) [9].
2008	Bitcoin: A chain of hash-based proof-of-work cryptocurrency [10] which popularized the application of Blockchain [11].

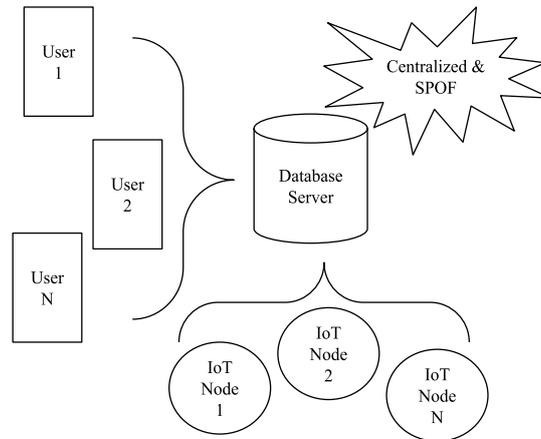


Fig. 1.1: A scenario of a traditional system handling requests from both IoT nodes and end-users which could lead to issues with centralization and SPOF.

ensuring sustainable growth. Similarly an image-based system for detecting vacant spaces in parking areas, combining edge density, closed contour density, and foreground/background pixel ratio for robust detection at low computational cost was proposed [18]. Later many researches have proposed different methods [19, 20, 21, 22] in order to address this parking space problem.

While IoT was bringing the physical things into the digital world, its architectural model as shown in Fig.1.1 had centralization leading to a SPOF(Single Point of Failure), Blockchain technology eliminated the

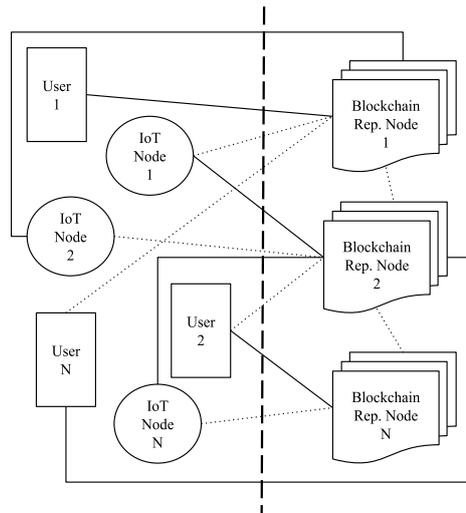


Fig. 2.1: A simplified decentralized architecture of using modern blockchains with IoT where requests from any entity are handled by various blockchain representative nodes and avoiding SPOF.

middleman for authoring or confirming a transaction. The integration of these two technologies was a need for a new breed of advancements.

**2. Literature Survey.** In 2015, an IoT electric business model based on the protocol of Bitcoin [23] was showcased which leverages blockchain technology and smart devices to revolutionize the energy sector. It enables seamless peer-to-peer energy trading, ensuring efficient utilization of renewable energy sources. IoT devices, like smart meters, monitor consumption, and autonomously interact with the blockchain, executing secure, transparent, and tamper-resistant transactions. Consumers can buy and sell excess electricity directly from one another, incentivizing green energy production and reducing dependency on traditional energy providers. The decentralized nature of the model as shown in Fig.2.1 eliminates intermediaries, reduces costs, and empowers individuals to participate actively in the sustainable energy economy. This innovation paves the way for a greener, more equitable energy future.

Securing smart cities using blockchain technology [24] offers a transformative approach to urban safety and efficiency. Blockchain's decentralized and immutable nature enhances data integrity, thwarting cyber threats and ensuring privacy in interconnected systems. Through tamper-resistant ledgers, critical infrastructure like transportation, energy, and healthcare can be safeguarded from unauthorized access and attacks. Smart contracts [25] can automate enforcement of regulations, streamlining governance and reducing corruption risks. Furthermore, blockchain's ability to facilitate secure peer-to-peer transactions fosters trust in smart city ecosystems, promoting seamless interactions between citizens, businesses, and government entities. Implementing blockchain-based solutions [26] [27] [28] can fortify smart cities against cyber threats [34], elevate citizen trust, and pave the way for a more sustainable and secure urban future.

Researchers started exploring the potential of blockchain for decentralized IoT [29] [30] architectures and discussed how blockchain could enhance trust, data integrity, and authentication in IoT networks. Blockchain-based solutions for IoT [31] device identity, access control, and data provenance gained attention leading to new consensus algorithms to handle the scalability challenges of combining IoT and blockchain. Integration of blockchain and IoT shifted towards energy management, supply chain, and smart city applications highlighting the potential benefits [32] of combining the blockchain's decentralized nature with IoT's data collection capabilities. Focus on blockchain-based solutions for securing IoT communications and enabling interoperability between devices explored the use of smart contracts to automate transactions and enforce rules [33] in IoT ecosystems. Continued research on blockchain's role in securing IoT data, improving trust, and addressing

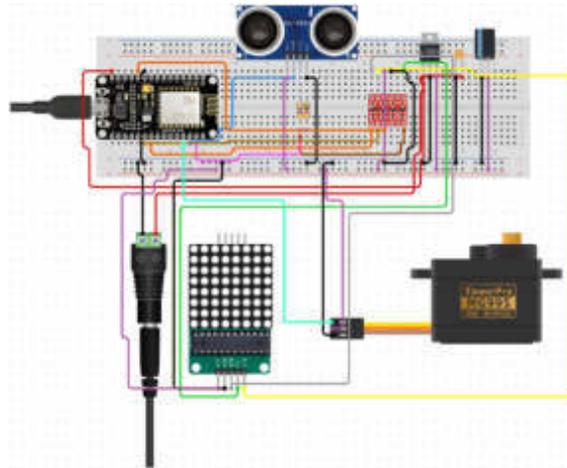


Fig. 3.1: Electronic design of the various components interconnected for a single IoT node in order to work with the proposed methodology at the entry/exit of the parking space

privacy concerns [35] discussed the challenges [36] [37] of integrating IoT devices with blockchain networks, such as resource constraints and high transaction costs.

Many advancements in parking space systems using computer vision & LoRa [38], AVI tags & RFIDs [39], transfer learning & deep learning [40] and Blockchain-IoT integration [41] [42] have bridged the gap between the needs of various solutions as the requirements change based on a problem being addressed. Hence, there is also a need to solve simple problems in Blockchain-IoT integration using economical and simple solutions which was taken up by this research and is proved by applying the methodology to a parking space counter.

**3. System Architecture.** The proposed system architecture as shown in Fig.3.2 integrates IoT and blockchain technologies to create a parking space counter with online viewing capability.

IoT devices Fig.3.1 are made up of HC-SR04 ultrasonic sensors that detect the approaching objects(vehicles) and MG995 10kg-cm metal gear servo motor that are used to control the motion of the entry/exit gates. A Wi-Fi communication module is attached with each IoT device that enables communication and access server APIs. The IoT devices are also equipped with P10 LED matrix display boards (interfaced via DMD connector) in order to present the current capacity of the parking lot.

Each IoT device can be equipped with a 40W solar panels and battery management systems in order to work standalone and also to reduce the energy consumption.

IoT devices installed at entry and exit of the parking lot detect available spaces by maintaining appropriate wallet balance securely stored on a blockchain network, ensuring transparency and immutability.

Users can access the available parking spaces online through a mobile app in real-time. Blockchain's decentralized nature ensures data integrity, while IoT enhances efficiency and accuracy in monitoring parking spaces, providing a seamless and transparent parking experience. It is to be noted that this research uses a blockchain which is actually meant for the sole purpose as a cryptocurrency in the use of developing a IoT application's solution.

**4. Implementation.** The IoT devices have to work inclined with the XNO blockchain to handle the parking lot spaces as shown in Fig.4.1. Initially the XNO entry wallet has to be funded with a balance that would be equal to the total number of parking spaces. For example if total parking spaces are equal to 25 the XNO entry wallet balance should be equal to 0.000025 and 25 is displayed on LED board at parking entry. In the beginning the XNO exit wallet balance is 0.

When IoT device detects car at entry gate then first the XNO entry wallet balance is checked, if balance is available then 1 unit i.e 0.000001 is transferred from XNO entry wallet to XNO exit wallet and the IoT device





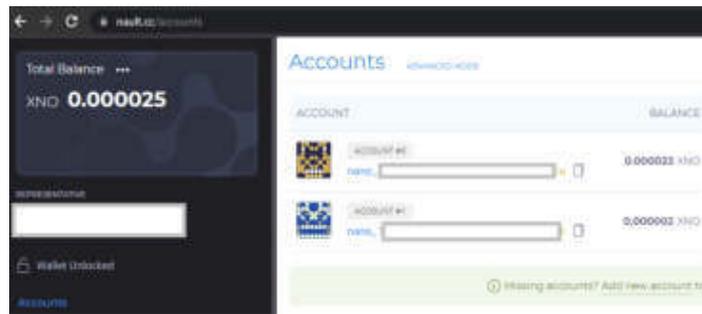


Fig. 4.2: An online wallet containing two XNO accounts: address of account#0 binded to parking entry and address of account#1 binded to the parking exit (source: nault.cc)

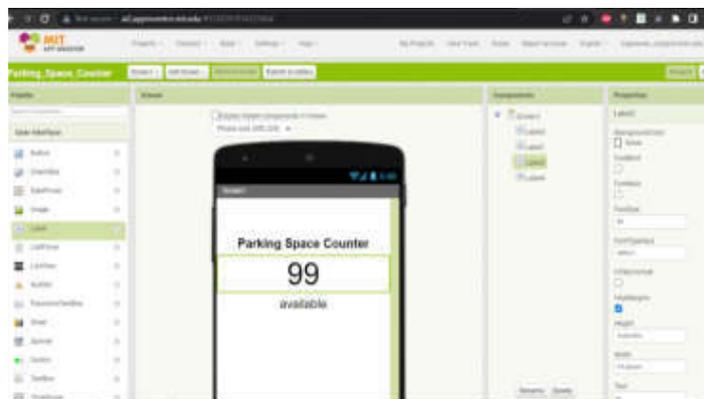


Fig. 4.3: Development scenario of the parking space counter app using MIT App Inventor (source: appinventor.mit.edu)

considered i.e. 1 IoT device at the entry gate of the parking lot and 1 IoT device at the exit gate of the parking lot. Note that XNO and its forks supports upto 30 decimal places and hence the cost change is very negligible when one parking lot contains huge number of parking spaces also.

An organization which tends to develop parking solutions based on this methodology will have to handle multiple number of parking lots. This is where XNO proves to be a viable solution as some of the other fee-less blockchain technologies tend to increase in price drastically as seen in Tab.5.3.

It was observed through this research that a cryptocurrency which is meant for money transfer can also be used to provide solutions in IoT. This solution was only possible because the block-lattice technology developed by Nano(XNO) made asset transfer fee-less and instant. Multiple parking entry and exit points can be extended with the same program without changes because the block-lattice is protected by double-spending acting as a process synchronization feature.

**6. Conclusion.** This research represents a significant advancement in parking management and efficiency by leveraging the Internet of Things (IoT) technology and incorporating the block-lattice mechanism, this system provides a reliable and real-time solution for monitoring and managing parking spaces in various urban and commercial settings.

The implementation of this system offers several key benefits. First and foremost, it enhances the overall

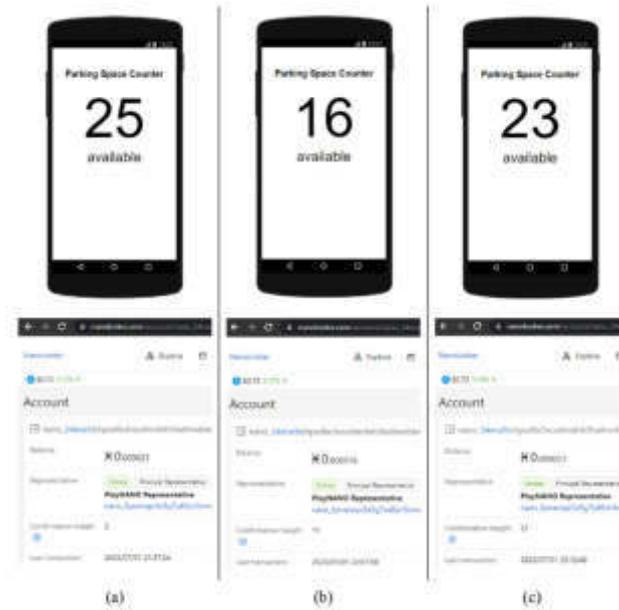


Fig. 5.1: Representation of parking space counter in accordance with wallet balance for various scenarios: (a)Initial Wallet balance is set equal to the total number of parking slots. 0.000025 Nano is represented on the phone display after conversion as 25microNano, (b)After some cars enter the parking lot the wallet balance decreases in proportion and updated on counter display at parking entry & app, (c)After some cars exit the parking lot the wallet balance increases in proportion. All updates are shown on display hoarding at parking entry & online using mobile app(photos above in first row).

Table 5.1: Attributes of various fee-less block-chain technologies that were considered for comparative study.

Block-chain	Max. TPS	Min. Token required (cost in USD)
EOS	8,000	0.0001 EOS (\$0.00006354)
IOTA	1,000	1 IOTA (\$0.158295)
XNO	1,800	0.000001 XNO (\$0.000000682)
BAN	XNO fork	0.01 BAN (\$0.00004091)

Table 5.2: Tokens required of the related block-chain and cost in USD for implementing the proposed methodology using various block-chain technologies for varying parking lot capacities ahving only 2 IoT devices; 1 at entry gate and 1 at exit gate.

BCT	for 25 spaces	for 50 spaces	for 100 spaces
EOS	0.0025 (\$0.0015939)	0.005 (\$0.0031879)	0.01 (\$0.0063758)
IOTA	1.000025 (\$0.17)	1.00005 (\$0.17)	1.0001 (\$0.17)
XNO	0.000001(\$0.0000007)	0.000001(\$0.0000007)	0.000001(\$0.0000007)
BAN	0.01 (\$0.00004)	0.01 (\$0.00004)	0.01 (\$0.00004)

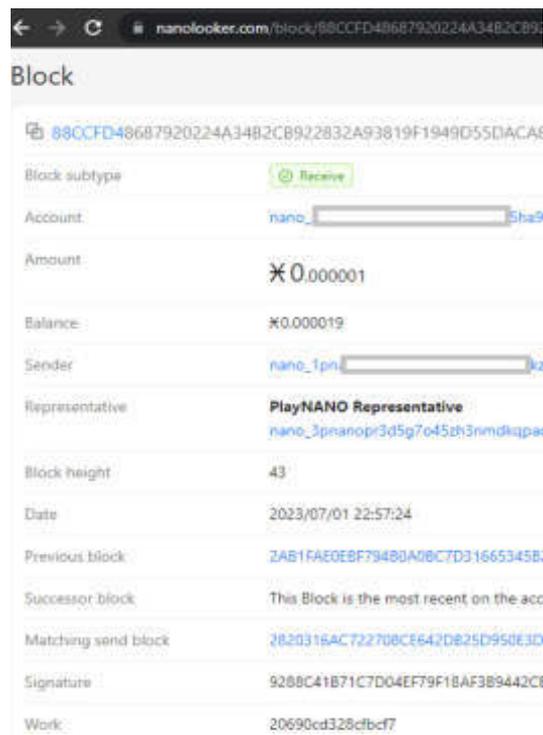


Fig. 5.2: Transaction block information which can be used to issue parking tickets as it contains timestamp and unique block hash (source: nanolooker.com)

Type	Account / Block	Amount	Date
Receive	nano_1pr... 88CCFD48687920224A3482CB922832A93819F1949D55DACA5	0.000001	2023/07/01 2 days ago
Send	nano_1pr... 2AB1FAE0EBF79480A08C7D31665345B	0.000001	2023/07/01 2 days ago
Receive	nano_1pr... 805AC125E3...	0.000001	2023/07/01 2 days ago
Receive	nano_1pr... 48E75E142F3...	0.000001	2023/07/01 2 days ago
Send	nano_1pr... 5282C0CCE...	0.000001	2023/07/01 2 days ago

Fig. 5.3: Transaction history of entry wallet where “Send” denotes a car entering the parking lot and “Receive” denotes a car exiting the parking slot (source: nanolooker.com)

parking experience for drivers by reducing the time spent searching for available spaces. This, in turn, leads to reduced traffic congestion and improved air quality in urban areas. Additionally, the system optimizes parking space utilization, maximizing revenue generation for parking lot operators and ensuring fair distribution of spaces among users.

Moreover, the integration of blockchain-based block-lattice technology ensures a secure and tamper-resistant

Table 5.3: Capital investments for various block-chain technologies if opted by organizations using the proposed methodology with multiple number of parking lots (with 100 spaces and 1 set of IoT devices) in different geographical locations.

BCT	for 10 lots	for 100 lots	for 1000 lots
EOS	0.1 (\$0.063758)	1 (\$0.63758)	10 (\$6.3758)
IOTA	10 (\$1.7)	100 (\$17)	1000 (\$170)
XNO	0.00001(\$0.000007)	0.0001(\$0.00007)	0.001(\$0.0007)
BAN	0.1 (\$0.0004)	1 (\$0.004)	10 (\$0.04)

data recording process in order to issue parking tickets. The decentralized nature of the block-lattice mechanism eliminates the risk of a single point of failure and enhances the system's resilience against potential cyber threats.

In conclusion, this innovative parking space counter system represents a step forward in transforming a cryptocurrency to solve traditional parking management by using an efficient, and user-friendly process. As IoT and blockchain technologies continue to evolve, we can expect even more simpler and integrated solutions to revolutionize urban mobility and shape smart cities of the future.

**Acknowledgments.** This work was funded by the Research & Development Cell(RDC), Maturi Venkata Subba Rao Engineering College, Osmania University, Hyderabad, Telangana, India. The authors thank NANO Foundation, Nault, SomeNano, NanoLooker and MIT App Inventor for related technology that was used in the implementation of this work.

#### REFERENCES

- [1] M. M. IRVINE, *Early digital computers at Bell Telephone Laboratories*, in IEEE Annals of the History of Computing, vol. 23, no. 3, pp. 22-42, July-Sept. 2001, doi: 10.1109/85.948904.
- [2] J. E. O'NEILL, *The role of ARPA in the development of the ARPANET, 1961-1972*, in IEEE Annals of the History of Computing, vol. 17, no. 4, pp. 76-81, Winter 1995, doi: 10.1109/85.477437
- [3] MIKE KAZAR, *CMU SCS Coke Machine*, 1982, <https://www.cs.cmu.edu/coke>
- [4] A. T. SHERMAN, F. JAVANI, H. ZHANG AND E. GOLASZEWSKI, *On the Origins and Variations of Blockchain Technologies*, in IEEE Security & Privacy, vol. 17, no. 1, pp. 72-77, Jan.-Feb. 2019, doi: 10.1109/MSEC.2019.2893730
- [5] J. ROMKEY, *Toast of the IoT: The 1990 Interop Internet Toaster*, in IEEE Consumer Electronics Magazine, vol. 6, no. 1, pp. 116-119, Jan. 2017, doi: 10.1109/MCE.2016.2614740
- [6] S. HABER, W.S. STORNETTA, *How to time-stamp a digital document*, in Journal of Cryptology, vol 3, no 2, pages 99-111, 1991
- [7] D. BAYER, S. HABER, W.S. STORNETTA, *Improving the efficiency and reliability of digital time-stamping*, in Sequences II: Methods in Communication, Security and Computer Science, pages 329-334, 1993
- [8] W. E. ZHANG ET AL., *The 10 Research Topics in the Internet of Things*, 2020 IEEE 6th International Conference on Collaboration and Internet Computing (CIC), Atlanta, GA, USA, 2020, pp. 34-43, doi: 10.1109/CIC50333.2020.00015
- [9] ISO/IEC 20922:2016 INFORMATION TECHNOLOGY, *Message Queuing Telemetry Transport (MQTT)*, v3.1.1, <https://www.iso.org/standard/69466.html>
- [10] NAKAMOTO, S., *Bitcoin: A Peer-to-Peer Electronic Cash System*, 2008, <https://bitcoin.org/bitcoin.pdf>
- [11] NATIONAL ARCHIVES AND RECORDS ADMINISTRATION, *Blockchain White Paper*, February 2019, <https://www.archives.gov/files/records-mgmt/policy/nara-blockchain-whitepaper.pdf>
- [12] S. KING AND S. NADAL, *Ppcoin: Peer-to-peer crypto-currency with proof-of-stake*, 2012, <https://peercoin.net/assets/paper/peercoin-paper.pdf>
- [13] C. LEMAHIEU, *Raiblocks distributed ledger network*, 2014, <https://docs.nano.org/whitepaper/english>
- [14] Y. RIBERO AND D. RAISSAR, *Dagcoin whitepaper*, 2015, <https://dagcoin.org>
- [15] S. POPOV, *The Tangle*, 2016, <https://www.iota.org/foundation/research-papers>
- [16] IETF RFC 7252, *Constrained Application Protocol (CoAP)*, 2014, <https://datatracker.ietf.org/doc/html/rfc7252>
- [17] J. ELLIOTT, H. JAYACHANDRAN, P. KUMAR AND K. METZER, *Campus shuttle: Design of a college campus parking and transportation system*, 2013 IEEE Systems and Information Engineering Design Symposium, Charlottesville, VA, USA, 2013, pp. 104-109, doi: 10.1109/SIEDS.2013.6549502
- [18] J. LIU, M. MOHANDS AND M. DERICHE, *A multi-classifier image based vacant parking detection system*, 2013 IEEE 20th International Conference on Electronics, Circuits, and Systems (ICECS), Abu Dhabi, United Arab Emirates, 2013, pp. 933-936, doi: 10.1109/ICECS.2013.6815565.
- [19] A. GRAZIOLI, M. PICONE, F. ZANICHELLI AND M. AMORETTI, *Collaborative Mobile Application and Advanced Services for Smart Parking*, 2013 IEEE 14th International Conference on Mobile Data Management, Milan, Italy, 2013, pp. 39-44,

- doi: 10.1109/MDM.2013.63.
- [20] M. DIXIT, C. SRIMATHI, R. DOSS, S. LOKE AND M. A. SALEEMDURAI, *Smart Parking with Computer Vision and IoT Technology*, 2020 43rd International Conference on Telecommunications and Signal Processing (TSP), Milan, Italy, 2020, pp. 170-174, doi: 10.1109/TSP49548.2020.9163467.
- [21] I. C. Y. M. N. SYAHRIL KAHARU, F. FARABI, F. I. HARIADI AND W. ADIJARTO, *Integrated Indoor Parking Information System*, 2021 International Symposium on Electronics and Smart Devices (ISESD), Bandung, Indonesia, 2021, pp. 1-5, doi: 10.1109/ISESD53023.2021.9501709.
- [22] N. V. V. RAVIPATI, V. R. N. MONDRETI, S. K. SATTI, C. S. GURUGUNTI AND V. N. S. L. JAKKAMPUDI, *Computer Vision-Based System for Locating and Counting Vacant Parking Lot*, 2022 IEEE International Conference on Data Science and Information System (ICDSIS), Hassan, India, 2022, pp. 1-6, doi: 10.1109/ICDSIS55133.2022.9915916.
- [23] Y. ZHANG AND J. WEN, *An IoT electric business model based on the protocol of bitcoin*, 2015 18th International Conference on Intelligence in Next Generation Networks, Paris, France, 2015, pp. 184-191, doi: 10.1109/ICIN.2015.7073830
- [24] K. BISWAS AND V. MUTHUKUMARASAMY, *Securing Smart Cities Using Blockchain Technology*, 2016 IEEE 18th International Conference on High Performance Computing and Communications; IEEE 14th International Conference on Smart City; IEEE 2nd International Conference on Data Science and Systems (HPCC / SmartCity / DSS), Sydney, NSW, Australia, 2016, pp. 1392-1393, doi: 10.1109/HPCC-SmartCity-DSS.2016.0198
- [25] K. CHRISTIDIS AND M. DEVETSIKIOTIS, *Blockchains and Smart Contracts for the Internet of Things*, in IEEE Access, vol. 4, pp. 2292-2303, 2016, doi: 10.1109/ACCESS.2016.2566339
- [26] M. SAMANIEGO AND R. DETERS, *Hosting Virtual IoT Resources on Edge-Hosts with Blockchain*, 2016 IEEE International Conference on Computer and Information Technology (CIT), Nadi, Fiji, 2016, pp. 116-119, doi: 10.1109/CIT.2016.71
- [27] M. SAMANIEGO, U. JAMSRANDORJ AND R. DETERS, *Blockchain as a Service for IoT*, 2016 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), Chengdu, China, 2016, pp. 433-436, doi: 10.1109/iThings-GreenCom-CPSCom-SmartData.2016.102
- [28] S. HUH, S. CHO AND S. KIM, *Managing IoT devices using blockchain platform*, 2017 19th International Conference on Advanced Communication Technology (ICACT), PyeongChang, Korea (South), 2017, pp. 464-467, doi: 10.23919/ICACT.2017.7890132.
- [29] S. HUH, S. CHO AND S. KIM, *Managing IoT devices using blockchain platform*, 2017 19th International Conference on Advanced Communication Technology (ICACT), PyeongChang, Korea (South), 2017, pp. 464-467, doi: 10.23919/ICACT.2017.7890132
- [30] Z. HUANG, X. SU, Y. ZHANG, C. SHI, H. ZHANG AND L. XIE, *A decentralized solution for IoT data trusted exchange based-on blockchain*, 2017 3rd IEEE International Conference on Computer and Communications (ICCC), Chengdu, China, 2017, pp. 1180-1184, doi: 10.1109/CompComm.2017.8322729
- [31] M. SINGH, A. SINGH AND S. KIM, *Blockchain: A game changer for securing IoT data*, 2018 IEEE 4th World Forum on Internet of Things (WF-IoT), Singapore, 2018, pp. 51-55, doi: 10.1109/WF-IoT.2018.8355182
- [32] G. WANG, Z. SHI, M. NIXON AND S. HAN, *ChainSplitter: Towards Blockchain-Based Industrial IoT Architecture for Supporting Hierarchical Storage*, 2019 IEEE International Conference on Blockchain (Blockchain), Atlanta, GA, USA, 2019, pp. 166-175, doi: 10.1109/Blockchain.2019.00030
- [33] IEEE STD 2144.1-2020, *IEEE Standard for Framework of Blockchain-based Internet of Things (IoT) Data Management*, 18 Jan. 2021, doi: 10.1109/IEEESTD.2021.9329260
- [34] KANTIPUDI, M. P., ALUVALU, R., VELAMURI, S., *An Intelligent Approach of Intrusion Detection in Mobile Crowd Sourcing Systems in the Context of IoT Based SMART City*, 2023, Smart Science, 11(1), 234-240.
- [35] ALUVALU, R., VN, S. K., THIRUMALAISAMY, M., BASHEER, S., SELVARAJAN, S., *Efficient data transmission on wireless communication through a privacy-enhanced blockchain process*, 2023, PeerJ Computer Science, 9, e1308.
- [36] Z. GONG-GUO AND Z. WAN, *Blockchain-based IoT security authentication system*, 2021 International Conference on Computer, Blockchain and Financial Development (CBFD), Nanjing, China, 2021, pp. 415-418, doi: 10.1109/CBFD52659.2021.00090
- [37] M. SHAIKH, C. SHIBU, E. ANGELES AND D. PAVITHRAN, *Data Storage in Blockchain Based Architectures for Internet of Things (IoT)*, 2021 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS), Toronto, ON, Canada, 2021, pp. 1-5, doi: 10.1109/IEMTRONICS52119.2021.9422573
- [38] B. C. MALLIKARJUN, G. ANANTHA KRISHNA, B. V. MAHANTESHA, K. T. MONISH AND E. KISHORE, *Parking Space Monitoring - A LoRa Based Approach*, 2023 11th International Conference on Internet of Everything, Microwave Engineering, Communication and Networks (IEMECON), Jaipur, India, 2023, pp. 1-5, doi: 10.1109/IEMECON56962.2023.10092339.
- [39] S. A. RATTI, N. PIRZADA, S. M. A. SHAH AND A. NAVEED, *Intelligent Car Parking System Using WSN*, 2023 Global Conference on Wireless and Optical Technologies (GCWOT), Malaga, Spain, 2023, pp. 1-9, doi: 10.1109/GCWOT57803.2023.10064656.
- [40] A. KHALFI AND M. GUERROUMI, *Transfer Learning with Image Data Augmentation for Parking Occupancy Detection*, 2023 International Conference on Advances in Electronics, Control and Communication Systems (ICAECCS), BLIDA, Algeria, 2023, pp. 1-4, doi: 10.1109/ICAECCS56710.2023.10105107.
- [41] S. SUN, H. TANG AND R. DU, *A Novel Blockchain-Based IoT Data Provenance Model*, 2022 2nd International Conference on Computer Science and Blockchain (CCSB), Wuhan, China, 2022, pp. 46-52, doi: 10.1109/CCSB58128.2022.00015
- [42] J. P. DE BRITO GONÇALVES, G. SPELTA, R. DA SILVA VILLAÇA AND R. L. GOMES, *IoT Data Storage on a Blockchain Using Smart Contracts and IPFS*, 2022 IEEE International Conference on Blockchain (Blockchain), Espoo, Finland, 2022, pp. 508-511, doi: 10.1109/Blockchain55522.2022.00078
- [43] NAULT, *A secure open source wallet for nano*, <https://github.com/Nault/Nault>
- [44] NANOVAULT, *A fully client-side signing wallet for sending and receiving Nano on your desktop or in your browser*,

<https://github.com/cronoh/nanovault>

- [45] NANO FOUNDATION, *Nano Node and Protocol Documentation*, <https://docs.nano.org>
- [46] SOMENANO, *A free-to-use Public Nano Node since 2021*, <https://node.somenano.com>
- [47] NANO FAUCETS, *Try nano (XNO) by getting a small amount for free*, <https://hub.nano.org/faucets>
- [48] NANO TOOLS, *NANO unit converter 1.2.6*, <https://nanoo.tools/unit-converter>
- [49] MIT, *MIT App Inventor*, <https://appinventor.mit.edu>
- [50] MIT APP INVENTOR, *Using Web APIs with JSON*, <http://ai2.appinventor.mit.edu/reference/other/json-web-apis.html>
- [51] CISCO NETACAD, *Cisco Packet Tracer*, <https://www.netacad.com/courses/packet-tracer>
- [52] NANOLOOKER, *NANO(XNO) Block Explorer*, <https://www.nanolooker.com>

*Edited by:* Kumar Abhishek

*Special issue on:* Machine Learning and Blockchain based Solution for Privacy and Access Control in IoT

*Received:* Jul 29, 2023

*Accepted:* Nov 30, 2023



## A STUDY OF BLOCKCHAIN AND MACHINE LEARNING-ENABLED IOT SECURITY IN TIME-DELAYED NEURAL NETWORK VOCAL PATTERN RECOGNITION TO IMPROVE WEB-BASED VOCAL TEACHING

KAIYI LONG\*

**Abstract.** With the development of information technology, online vocal teaching is becoming more and more popular, but the sound quality of teaching is also becoming more and more demanding. As online vocal instruction becomes more popular, the need for high-quality sound in these digital environments becomes more critical. This research tackles the problem of improving sound quality in real-time vocal teaching by integrating advanced technologies such as Blockchain and Machine Learning within the Internet of Things (IoT) security framework. We created a vocal recognition model using Time-Delay Neural Network (TDNN) and improved it with Generated Feature Vector (GFV). This integration yields a strong GTDNN vocal recognition system that is specifically designed to secure and optimize web-based vocal teaching. Our experiments show that GTDNN outperforms traditional TDNN and i-vector methods in feature vector extraction, adapting well to different speech environments. In various speech settings, GTDNN's Error Rates (EERs) are impressively low at 11.3%, 12.0%, 4.9%, 6.2%, and 6.1%, indicating superior performance over comparison models. GTDNN has an EER of 9.6% for short-duration speech and 2.3% for long-duration speech. Furthermore, the GTDNN system achieves an overall pass rate of 94% for target speech and an impressive rejection rate for non-target speech, ensuring high accuracy in a variety of speech environments.

**Key words:** TDNN; vocal recognition; network vocal teaching; sound quality fidelity

**1. Introduction.** At present, the application of network technology in teaching is one of the forms of modernization of education and teaching in China. This form is also reflected in vocal music teaching, which makes it possible to change from traditional offline teaching to online teaching, which greatly increases the flexibility of teaching and the richness of teaching forms. In the process of online vocal teaching, there are inevitably background sounds and other people's voices mixed in the teacher's teaching voice, so online sound quality fidelity technology is needed to improve the sound quality. Voice recognition can identify and distinguish teachers from other voices by extracting their vocal features, thus achieving sound quality fidelity [3, 19]. Voice is a kind of information carrier for communication between people, and voice contains a variety of information, for example, in daily life, people can be distinguished by their voice alone. This is due to the fact that the vocal pattern of each person's voice is different. Therefore, vocal recognition technology can be applied to the task of recognizing speakers. In short, it is the extraction of features in a speaker's voice that characterize the speaker's voice and the analysis and classification of these identity features to distinguish target and non-target speakers [5, 16].

At its core, blockchain is a distributed database that keeps a constantly growing list of records, known as blocks, that are linked and secured using cryptography. Each block contains a cryptographic hash of the previous block, as well as a timestamp and transaction data, resulting in an unchangeable chain of records. This ensures that once a record has been added to the chain, it cannot be changed retroactively without affecting all subsequent blocks, which requires network majority consensus. Because blockchain is decentralized, there is no need for a central authority or intermediary, making it naturally resistant to fraud and corruption. Its applications span a wide range of industries, including finance, healthcare, supply chain management, and others, where it promises increased transparency, security, and efficiency. Vocal recognition has a wide range of application scenarios as scientific research continues to deepen [6]. Especially in recent years, with the continuous development of machine learning theories, machine learning techniques have made new breakthroughs in natural language processing and data mining. In recent years, with the significant breakthrough of deep learning

---

\*The School of Film and Television Art, Hunan Mass Media Vocational and Technical College Changsha, 410100, China ([kaiyilong@outlook.com](mailto:kaiyilong@outlook.com))

theory in speech recognition, more and more researchers have started to use it for voiceprint recognition. There are numerous approaches to use deep learning techniques for voice recognition, but most of the current systems using deep learning still embed neural network models into the identity vector (i-vector) recognition benchmark framework based on Probabilistic Linear Discriminant Analysis (PLDA) models model at [6]. However, the recognition rate of short-time speech is low, and it is easy to be disturbed by noise. Time-Delay Neural Network (TDNN) is a model using multi-layer convolutional neural network, which can carry out convolution operations on both time axis and frequency axis, and has stronger robustness. In order to improve the recognition rate and stability of the voicing system, and strengthen the robustness against noise, this research will build a vocal recognition method based on TDNN to accurately identify the teacher's voice in online lessons and improve the quality of vocal teaching. The combination of these technologies — blockchain for security, machine learning for pattern recognition, and IoT for connectivity — has the potential to significantly improve the dependability and efficiency of web-based vocal teaching platforms. This study aims to improve the overall experience and outcomes of digital vocal education by ensuring secure data exchanges and accurate vocal recognition.

**2. Literature Review.** In the field of acoustics teaching, researchers have produced results on the reform of teaching mode. Yang h analyzes the quality of multimedia teaching in vocal music class by using elite teaching optimization algorithm to improve the level of acoustic teaching. The main point of view is the teacher's teaching ideas, the level of information-based teaching design, the level of new teaching model, teaching effects and other factors that affect the quality of teaching. The results show that the algorithm is effective and can be used to evaluate the quality of vocal music multimedia teaching [20]. Fu L thinks that the vocal music teaching is changing from teacher-centered to student-centered. The teaching emphasizes the development of students and the change of academic conditions, and pays more attention to the combination of theory and practice. Fu L analyzed the actual situation of vocal music teaching to find out the gap between reality and theory, and thus formulated a more scientific teaching method to improve the teaching model. The study provides a theoretical teaching plan for improving the quality of acoustic teaching and the acoustic level of students [3]. [16], in this paper, a method of system resource allocation based on power iteration is proposed to optimize the resource allocation of vocal music teaching system. The method takes the throughput of the unloading process as the objective function, and achieves the optimal allocation of normal power through iterative optimization. At the same time, a heterogeneous network based on edge server is proposed to improve the low energy efficiency and resource utilization of edge server. The experimental results show that the method is effective and practical.

Academic research on TDNN has produced many results. hu S et al. used TDNN optimized by neural structure search (NAS) based techniques for speech recognition tasks. the NAS was used to automatically learn two hyperparameters of factorized TDNN, namely left-right splicing and contextual offset, and the linear projection dimension of each hidden layer, allowing TDNN to perform in different systems through parameter sharing effective search. Experimental results show that its word error rate is large and model size is greatly reduced, and speech recognition performance is improved [9]. [1] used TDNN to predict the active power demand on a P4 bus in President Prudente. Experimental results demonstrated its validity [11]. [21] used TDNN as a facial expression classifier for an intelligent robot to establish command laws by analyzing and recognizing facial expressions to translate expressions into robot-recognizable language. Experimental results verified its effectiveness and improved the efficiency of recognition. [2] used TDNN and normalization methods for optimizing an automatic speech recognition system. The experimental results showed that the recognition error rate of the optimized system was greatly reduced and its acoustic and language model recording speed was significantly accelerated.

The paper [11] investigated the performance of deep neural networks, convolutional neural networks, temporal convolutional neural networks, and TDNN for English dialect classification. The results showed that TDNN and ECAPA-TDNN classifiers capture a wider temporal context, further improving the performance of the classification models. TDNN improved the performance of SPEC-STFT and SPEC-SFF by 2.8% and 1.4%, respectively. [15] applied focal time delay neural network (FTDNN) to ECG lead set accurate reconstruction. The experimental results showed that the FTDNN method reconstructed leads with correlation values between 0.8609 and 0.9678 and root mean square error values between 123  $\mu V$  and 245  $\mu V$  in all cases except for individual subgroups, outperforming other methods and improving the accuracy of ECG leads. [12] applied a modified TDNN to monitor the flow of granular solid material in oil pipes. The method can capture dynamic

acoustic signals. Experimental results showed that its normalized root mean square error for monitoring solid flow rate, solid concentration, pipeline pressure drop and gas velocity were 0.18, 0.17, 0.20 and 0.16, respectively. Its performance provides a simple and reliable real-time monitoring system due to the artificial neural network model. [14] used TDNN to build a three-level coding system to decode the signal when rats act. The study first recorded and processed M1 signals from three behavioral types of rats: walking, standing and head shaking, and then divided the signals according to response time points and analyzed them as independent components to extract signal features. Finally, the study finds 16 representative sample signals by dynamic dimension increase algorithm to input a three-level TDNN for training. The results showed that the recognition accuracy of TDNN for these three actions was 51.4%, 80.0% and 54.3%, respectively, which showed that the three-stage TDNN coding model could explain the M1 brain signals of free-motor animals [12].

To sum up, TDNN has rich application results in classification tasks, signal recognition and speech recognition, but few studies have applied it to voice print recognition and vocal music teaching tasks. TDNN can be used to improve the recognition accuracy rate and stability in noisy environment by its strong robustness. This research will apply TDNN in the field of vocal recognition, design a vocal recognition system based on TDNN that can be used for web-based vocal teaching, and provide a better performance method for online lesson sound quality fidelity.

### 3. Vocal Sound Quality Fidelity Method Based On Voiceprint Recognition With Time Delay Neural Network.

**3.1. Design of voice-print recognition model based on time-delay neural network.** Vocal recognition refers to the process of extracting features that can represent the identity of a speaker from one or more segments of speech and comparing them with the identity features extracted from the speech of a known speaker to achieve confirmation of the identity of the speaker. Its main processes are specifically as follows: preprocessing of speech signals, feature extraction of speech signals, training and building speaker models, and scoring the unknown speaker against the feature parameters of the known speaker [10]. Vocal recognition feature extraction is a key step in vocal recognition, and it is necessary to obtain the long time dependent nature of speech to obtain a time invariant acoustic model of this nature. TDNN can obtain the long time dependent nature of speech, and can extract the relationship between each feature value related to the time series, and take into account the characteristic factor of the change of feature value due to time change [13]. Therefore, TDNN is chosen for this study to perform vocal feature extraction. TDNN is a feedforward neural network that is computed using interconnected layers composed of many nodes, and its basic structure is shown in Figure 3.1. In the figure,  $a_i, a_j$  is the input vector of  $i$  dimension and  $j$  dimension,  $w$  is the connection weight,  $s_1, s_t$  is the delay vector, and  $t$  is the delay time. When  $t = 2$  is used, it means that the current time and the vectors of the two preceding and following frames are combined together and fed into the network. The core idea of TDNN is weight sharing, i.e., the weights of the same parts are the same. Specifically, when  $t = 2$ , there will be 3 weights in the neural network, which are given to the current time frame and the two frames before and after the current time.

The activation function can make the neural network have stronger classification ability. ReLU function and Softmax function are both commonly used activation functions. Among them, the ReLU function can effectively overcome the problem of gradient disappearance and can transfer the error better than the Sigmoid function. Therefore, the ReLU function is chosen as the activation function of TDNN in this study, and its formula is shown in equation 3.1.

$$\text{Max}(0, x) = \begin{cases} 0 & \text{if } x \leq 0 \\ x & \text{if } x \geq 0 \end{cases} \quad (3.1)$$

For the output classification function of TDNN, the Softmax classification function was chosen for the study, and its formula is given in equation 3.2.

$$y_i = \frac{\exp(x_i)}{\sum \exp(x_i)} \quad (3.2)$$

To design a vocal pattern extraction model based on TDNN, the model can be divided into two modules: speaker identity feature vector extraction and back-end scoring module. The former can utilize Mel-Frequency

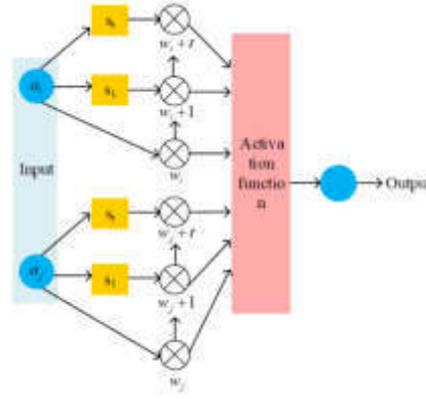


Fig. 3.1: Basic structure diagram of TDNN network

Cepstral Coefficients (MFCCs) features of fixed speech signals as input features for TDNN. each delay layer in TDNN reads the information features in the speaker's speech while expanding the context of the frame, which facilitates capturing the speaker's identity information. The delay layer is followed by a pooling layer, where TDNN collapses the features along the time axis by computing the mean and standard deviation statistics of the features to obtain a feature vector containing global speaker information. In order to get the speaker embedding correct, the pooling layer is followed by a fully connected layer. The back-end scoring module can be chosen from the PLDA model. The scoring principle of the PLDA model is to assume that the training data speech consists of speaker's speech, where each speaker has different segments of his or her own speech. Then, the  $b$  speech of the  $a$  speaker is defined as  $x_{ab}$ . Then, the generative model of can be defined as equation 3.3 based on the factor analysis.

$$x_{ab} = \mu + Fh_a + Gw_{ab} + \epsilon_{ab} \quad (3.3)$$

In equation 3.3,  $\mu$  is the numerical mean,  $F$  and  $G$  are two spatial feature matrices, represents the feature representation of in speaker space, and is the noise covariance. In the recognition scoring stage, if the likelihood of two voices having the same  $h_a$  features is greater, then the two voices belong more definitely to the same speaker. The degree of likelihood is used to calculate the score by the log-likelihood ratio, which is given in equation 3.4.

$$S = \log \left( \frac{p(a_1, a_2 | T_1)}{p(a_1 | T_2)p(a_2 | T_2)} \right) \quad (3.4)$$

In equation 3.4,  $S$  is the log-likelihood ratio.  $T_1$  and  $T_2$  are two hypothetical events, i.e., the vector  $a_1, a_2$  belongs or does not belong with a speaker.  $p$  represents the probability of the event. the PLDA model is based on two basic assumptions: first, that the channel influence is independent of the speaker; and second, that the total variability factor of the speaker cannot fluctuate excessively. However, these two assumptions are not fully valid. Therefore, the study improves it by eliminating the process of solving the spatial feature matrix  $G$  in equation 3.3, without reference to the differences between different speech sounds of the same speaker, to obtain a simplified PLDA model. The equation of the improved model is given in Equation 3.5.

$$x_{ab} = \mu + Fh_a + \epsilon_{ab} \quad (3.5)$$

The training phase of the neural network requires the selection of a suitable loss function, and this time the softmax cross-entropy function is chosen as the loss function, the formula of which is given in equation 3.6.

$$E = \sum_{n=1}^N \sum_{i=1}^n y_i \log \left( \frac{e^{z_i}}{\sum_k e^{z_k}} \right) \quad (3.6)$$

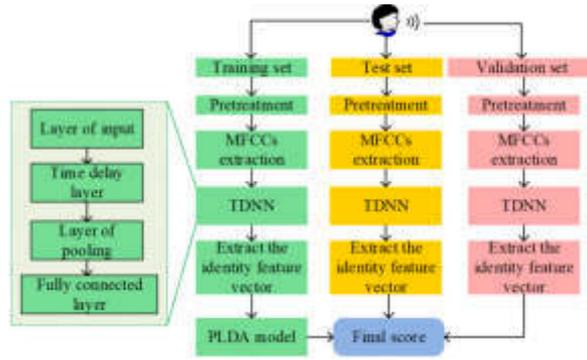


Fig. 3.2: Process of voice print recognition model based on TDNN

In equation 3.6,  $n$  is the number of data samples, and each data  $x_n$  contains  $m$  frames of sound feature data. If the speaker's label is now  $i$ , then  $y_i$  is 1; otherwise, it is 0.  $z_i$  denotes the input value of the first  $i$  node, and  $z_k$  denotes the input values of other nodes. Combined with the above design, the obtained time-delay neural network-based voice recognition model flow is shown in Figure 3.2. In the training phase, the speech data of the training set is inputted into the time-delay neural network with the extracted MFCCs after speech preprocessing and Mel cepstral coefficient extraction, and the output data of the softmax layer is used for the training of the PLDA back-end. In the testing phase, it is also necessary to first perform the speech preprocessing operation on the test set data to extract the MFCCs in the voice and input the MFCCs into the time-delay neural network, unlike the test phase, the feature vectors extracted by the neural network are presented after the hidden layer7 of the time-delay neural network and used directly for scoring the PLDA model.

**3.2. Model optimization and identification system construction.** TDNN considers multi-frame features, which enhances the complexity of the computation. Therefore, the study introduces the Generated feature vector (GFV) to improve it. The basic principle of the Generated feature vector is to perform correlation analysis between the Identity authentication vector (i-vector) and the extracted vectors of the TDNN model using Canonical correlation analysis (CCA) to generate a new feature vector. CCA allows the two feature vectors to learn from each other. To achieve voice recognition, it is necessary to find an invisible vector space in which each speaker is a point, i.e., a set of basis vectors of this vector space can be used to represent this speaker [18]. The identity feature vector i-vector is the vector belonging to this vector space that can characterize the speaker. In the registration and testing phases, the i-vector will be excluded from the system and only the feature vector extracted by the TDNN network will be considered and linearly transformed using the transformation matrix obtained by typical association analysis methods. This paper utilizes this transformed output as a generative feature vector, which captures some properties of the identity feature vector extracted by the i-vector model during the transformation process. The steps are shown in Figure 3.3. After inputting the speech, the matrix based on the identity vector  $W_a$  and the matrix extracted by TDNN  $W_b$  are transformed by using CCA learning, and then this transformation matrix is used for generative feature vector extraction.

The matrix  $W_a$  and the matrix  $W_b$  need to satisfy the constraints of equation 3.7.

$$\max_{W_a, W_b} \text{corr}(W_a \phi_a, W_b \phi_b) \quad (3.7)$$

In Eq. 3.7,  $\phi_a$  is the feature vector extracted from the i-vector and  $\phi_b$  is the feature vector extracted from the TDNN of the same speech. The CCA transform can transfer the information from the i-vector model to the TDNN model and vice versa. The expression of the generative feature vector is given in equation 3.8.

$$\phi_g = W_b \phi_b \quad (3.8)$$

In equation 3.8,  $\phi_g$  is the generative feature vector and also the feature vector extracted by TDNN with i-vector features. i-vector and TDNN networks extract feature vectors that are both zero-centered, but the

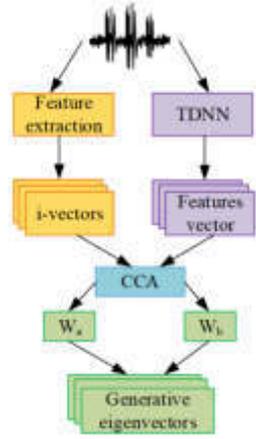


Fig. 3.3: Schematic diagram of GFV principle

dimensionality of the two is different. Applying CCA can maximize the correlation between the two. the principle of CCA is that if there are two random vectors  $X = \{x_1, x_2, \dots\}^T$  and  $Y = \{y_1, y_2, \dots\}^T$ , CCA can redefine the new variables  $M$  and  $N$  by a linear combination of  $X$  and  $Y$ .  $M$  and  $N$  see equation 3.9.

$$\begin{cases} M = a^T X \\ N = b^T Y \end{cases} \quad (3.9)$$

In Eq. 3.9,  $a^T$  and  $b^T$  are the two parameters that CCA needs to find to maximize the correlation coefficient of the two vectors. The expression of the correlation coefficient  $r$  is given in Eq. 3.10

$$r = \text{corr} \langle a^T X, b^T Y \rangle = \frac{E(a^T X b Y^T)}{\sqrt{E(a^T X a X^T)} \sqrt{E(b^T Y b Y^T)}} \quad (3.10)$$

In equation 3.10,  $E$  denotes the unit matrix. The constraints of this equation are shown in equation 3.11.

$$\begin{cases} a^T \sum_X a = 1 \\ b^T \sum_Y b = 1 \end{cases} \quad (3.11)$$

In equation 3.11,  $\sum_x$  and  $\sum_T$  are the covariances of the matrices.  $\sum_x$  The expressions are given in Eq. 3.12.

$$\sum_X = E(X^T X) \quad (3.12)$$

For  $\sum_Y$  the expression of is shown in equation 3.13.

$$\sum_Y = E(Y^T Y) \quad (3.13)$$

The relevant parameters for the maximization are given in equation 3.14.

$$r = a^T \sum_{XY} b = a^T E(XY^T) b \quad (3.14)$$

In CCA, the aim is to find mutually orthogonal pairs of the maximum correlated linear combinations of variables in  $X$  and  $Y$ . In the study,  $X$  and  $Y$  refer to the corresponding i-vector and TDNN network extracted

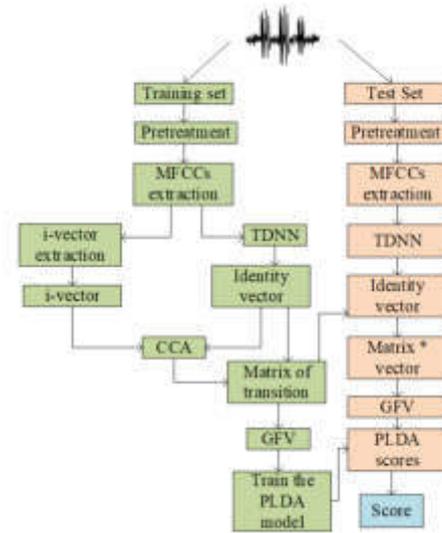


Fig. 3.4: Architecture of GTNN-based voice recognition system

feature vectors of the same speech segment, respectively. After introducing GFV into the voicing recognition model, a voicing recognition system based on time-delay neural network is designed. The system is developed in Python language under Linux environment and uses Advanced RISC Machine (ARM) processor for data processing. The system first obtains the audio from peripherals and stores it in the memory unit, then drives the AXI-DMA unit to transmit the data to the on-chip processor after processing by the processor, then uses the ARM processor to control the CNN hardware accelerator to read the features and weights, and finally sends the structure back to the on-chip memory and then sends it back to the memory via AXI\_DMA. Process the output further. The processing flow of the system for voiceprint recognition is shown in Figure 3.4

As shown in Figure 3.4, the system only has the extraction process for i-vectors in the training phase, but not in the test set, which reduces the amount of operations in the system. In the training phase, the i-vector feature extraction module is introduced mainly to calculate the transformation matrix  $W_a$ , which is used for generative feature vector generation. In the training phase of the model, the speech in the training set is first preprocessed to extract the acoustic parameters MFCCs in the speech, and then the i-vector of the speech is extracted and the TDNN network is used to extract the feature vectors of the speech, and two transformation matrices  $W_a$  and  $W_b$  are found using CCA to make the maximum correlation between the two vectors. One of the matrices  $W_b$  will be used to generate the test phase generative vectors. In the test phase, the speech is also preprocessed first to extract the acoustic parameters MFCCs in the speech, and the MFCCs extracted from the speech are input to the TDNN network as the input layer of the neural network. The TDNN network extracts the feature vectors that represent the identity of the speaker, and multiplies the feature vectors with the transformation matrix  $W_b$  calculated in the training phase to calculate the generative feature vectors in the speech. Finally, the generative feature vectors are input to the back-end scoring PLDA model for scoring, and the scoring results are used to determine whether the speech belongs to the target speaker to achieve the purpose of voice recognition.

The use of TDNN for vocal feature extraction is a significant step forward in vocal recognition technology. TDNN can extract time-invariant acoustic models, which are critical for identifying and verifying speaker identities with high accuracy, by capturing the long time-dependent nature of speech.

#### 4. Analysis of Experimental Results of GTNN-based Voice Recognition System.

**4.1. Experimental environment and parameter determination.** The environment for this experiment is shown below: the operating system is Linux Ubuntu18.0464, the GPU is NVIDIA TeslaP100, the CPU

Table 4.1: Basic parameters of sound library

Voice Library	Total number of speakers	Number of men	Number of women	Total number of voices	Number of voices per capita	Recording Environment	Total time/h
TIMIT	630	440	192	6300	10	Pure	3.96
Aishell	402	188	210	141600	355	Pure	177
VoxCeleb	1250	690	566	153520	120	Contains noise	353

Table 4.2: Experimental results of frame level layer delay configuration

Program Number	Delay layer number						loss	EER
	1	2	3	4	5	6		
1	{T, T+2}	{T-1, +3}	{T-2, T+4}	{T-3, T+6}	{T-4, T+6}	{T-5, T+7}	0.33	0.123
2	{T, T+2}	{T, T+2}	{T-1, T+3}	{T-1, T+3}	{T-2, T+4}	{T-2, T+4}	0.30	0.164
3	{T, T+2}	{T-1, T+3}	{T-1, T+3}	{T-2, T+4}	{T-2, T+4}	{T-2, T+4}	0.22	0.160
4	{T, T+2}	{T-1, T+3}	{T-1, T+3}	{T-2, T+4}	{T-2, T+4}	{T+1}	0.19	0.131
5	{T-1, T+3}	{T-1, T+3}	{T-2, T+4}	{T-2, T+4}	{T-3, T+5}	{T-3, T+5}	0.15	0.092
6	{T-1, T+3}	{T-1, T+3}	{T-2, T+4}	{T-3, T+5}	{T+1}	{T, T+1}	0.13	0.072

is Intel(R) Xeon(R) CPU is E5-2697V424, and the running memory is 96G. the software is developed using Python. The speech recognition data are obtained from existing publicly available voice databases, including the English dataset TIMIT voice library, VoxCeleb voice library and Chinese voice library Aishell, and the noise database is selected from MUSAN. the MUSAN dataset includes 929 various noise files with a sampling frequency of 16 KHz. all audio files are in WAV format. The basic parameters of other speech libraries are shown in Table 4.1. the TIMIT speech database has a sampling frequency of 16 kHz and includes a total of 630 target speakers, each recording 10 voices. voxCeleb is an audiovisual dataset covering video and audio data, and only the speech data are taken in the storage experiment. the Aishell speech library is a Chinese speech library covering various types of factual information.

The data of the speech library is divided into training set, registration set and test set, the training set is used to train TDNN and PLDA models, the registration set is used as the reference data set for recognizing voice patterns, and the test set is used to test the system performance. The Equal Error Rate (EER) is used as the main evaluation index of recognition effect. When the EER is larger, it means that the recognition system is less effective, and when the EER is smaller, it means that the recognition system is more effective.

Before starting the experiments, preliminary tests are needed to determine the optimal parameters of the TDNN [23, 24, 25]. The parameters to be determined include the delay time of each layer of the TDNN and the position of the extracted feature vectors. The first 6 layers of the TDNN model are the delay layer. In order to verify the effect of the different delay time of each layer on the experimental results and to find the optimal selection scheme, 6 schemes are tested in this experiment and the EER value as well as the loss value of each scheme are calculated. The experimental results of the delay configuration of the frame-level layers obtained are shown in Table 4.2. From the above table it is easy to see that the recognition effect of the delayed layer delay scheme selection of scheme 6 is optimal, so the delay parameters of each layer of the hidden layer of the delayed neural network are the same as in experiment 6.

In order to extract feature vectors with good distinguishability and represent the identity of the speaker, the effect of extracting feature vectors after fully connected layer 1 and fully connected layer 2 is tested for voice recognition, respectively. The dimensionality of the feature vector corresponds to the number of nodes in the fully connected layer. In order to extract the appropriate dimensionality of the feature vector, the effect of the number of nodes in the fully connected layer on the recognition results was tested with 256, 350 and 512 nodes, respectively. The VoxCeleb speech library was used for both testing and training, and the scoring backend was a PLDA model. The final model recognition results EER are shown in Table 4.3 below: from Table 4.3, it can

Table 4.3: Experimental results of feature vector extraction location

/	Feature vector extraction location	
	Fully connected layer 1	Fully connected layer 2
256	0.819	0.820
352	0.800	0.770
512	0.700	0.728

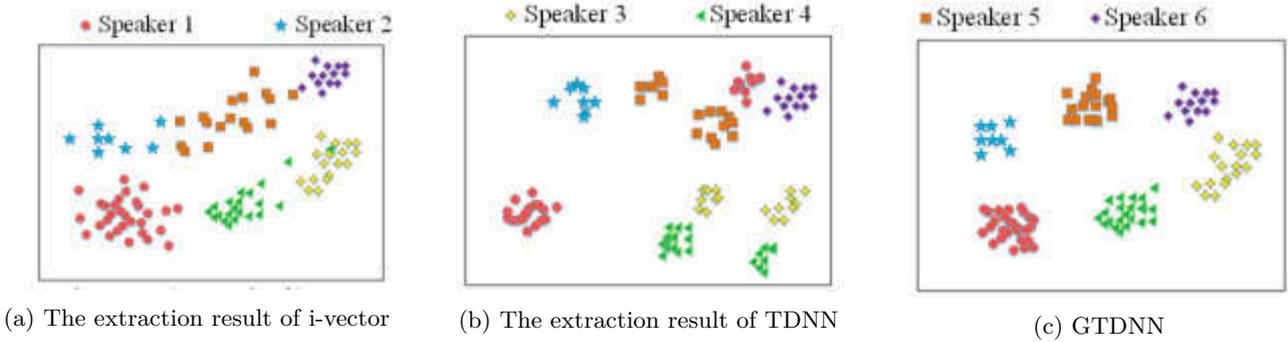


Fig. 4.2: Classification effect of extracted feature vectors

be seen that the best recognition results are obtained when the number of nodes is 512 and the feature vectors are extracted after the fully connected layer 1.

Therefore, the TDNN's delay parameter is chosen as option 6, and the feature vector extraction position is fully connected layer 1 when the number of nodes is set to 512.

**4.2. Effect analysis of feature vector classification.** To be able to visualize the performance progress of GTDNN in extracting feature vectors effect, the speaker model is visualized using t-sne (Distributed Stochastic Neighbor Embedding) high-dimensional data visualization technique to examine the speaker's identity vector feature extraction effect. Six speakers are randomly selected from the database, each with at least thirty speech items, and the feature vectors in the speech are extracted using three methods, i-vector, TDNN, and GTDNN, respectively. Then, to verify the effectiveness of the system, GTDNN is compared with TDNN, the continuous Markov model with real-time embedding from the literature (He and Dong 2020), and the hybrid acoustic model based on PDP coding from the literature [22]. In order to detect the robustness of the recognition system to noise and to compare the recognition effect of different systems, five different voice libraries were set up in this experiment, namely TIMIT voice library, TIMIT voice library with noise, Aishell voice library, Aishell voice library with noise and VoxCeleb voice library. t-sne visualization results are shown in Figure 4.2. The feature vectors extracted by GTDNN have strong differentiability, and the classification effect of its extracted feature vectors is significantly better than the classification effect of the feature vectors extracted by TDNN and i-vector based methods.

**4.3. The recognition effect analysis of the system.** The EER results for different environments are shown in Figure 6, where nTIMIT and nAishell denote the TIMIT speech bank with noise and the Aishell speech bank with noise. All systems perform better in the pure speech environment than in the noisy environment. In the overall performance, GTNN is the best recognition among several models. the continuous Markov model with GTNN and real-time embedding and the hybrid acoustic model based on PDP coding outperform the traditional TDNN model. Taking the VoxCeleb speech library, which is closest to the real environment, as an example, the EER of GTDNN is 6.1% in the test environment of the VoxCeleb speech library, which is 2.5% lower than the traditional TDNN model. In other speech libraries, their EERs are 11.3%, 12.0%, 4.9% and

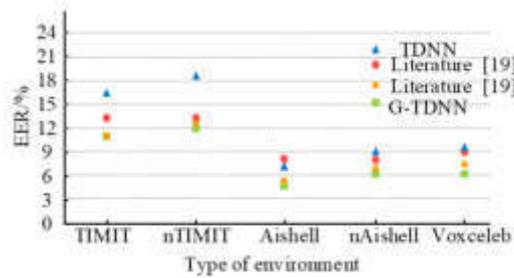


Fig. 4.3: EER results in different environments

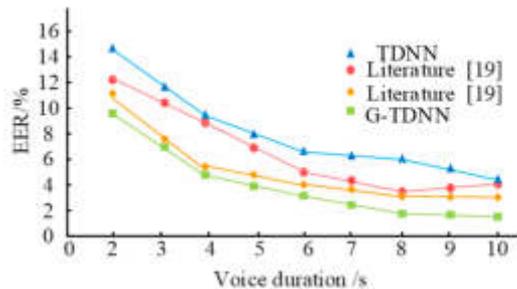


Fig. 4.4: Recognition results with different speech durations

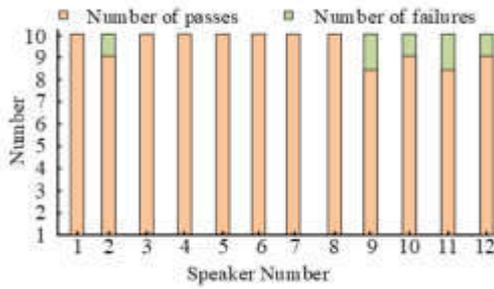
6.2%, which are all at a lower level.

The recognition EER results for different speech durations are shown in Figure 4.4. Overall, it seems that the EERs of all four systems are decreasing as the duration increases, and the recognition effect of each model is getting better. the TDNN model approach is worse for short-time speech of 2 s, and the recognition EER is the largest. the EER of GTDNN is the smallest among the four models for both short-time and long-time speech, with EER value is the largest at 9.6% and the smallest at 2.3%. The EERs of the continuous Markov model with real-time embedding and the hybrid acoustic model based on PDP coding are always in between those of TDNN and GTDNN. It can be seen that GTDNN has better recognition performance regardless of the speech environment and speech duration.

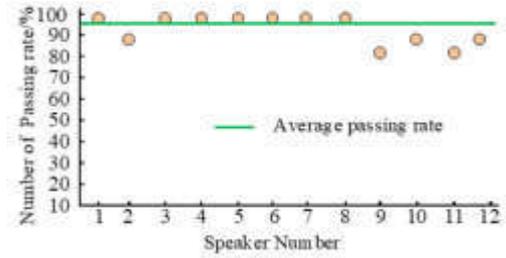
The results of the correct speech recognition pass rate of GTDNN are shown in Figure 4.6, from which it can be seen that the recognition pass rate of most people is above 90%, and the overall recognition pass rate can reach 94%, which can meet the needs of daily human-computer interaction. It can be seen that GTDNN has a high recognition rate and recognition stability.

The rejection rate of detecting non-target speech is shown in Table 4.4, and it can be seen from Table 4 that the rejection rate of GTDNN is relatively high, with an overall rejection rate of 94% and a rejection rate of 90% for most non-target speech, and the stability is also relatively strong, but there is still a risk of misidentification, and in practical applications one can choose to integrate other recognition methods to further improve the recognition accuracy. In summary, the application test using the system proposed in the study has a good recognition performance and the recognition effect is also stable, but there is still a partial false pass, and the evaluation threshold after scoring can be adjusted for different application scenarios in practical applications. In the system with the increase of the judging threshold, the false rejection rate of the system will increase to, while the smaller the judging threshold, the passing rate of correct speech will increase and the false acceptance rate of the system will decrease.

**5. Conclusion.** In the information age, online lessons have become a new mode of teaching in all majors, and this is also true for vocal majors. However, vocal music teaching requires a higher quality of sound for



(a) Number of passes and failures



(b) Failure Rate

Fig. 4.6: Recognition pass rate results

Table 4.4: Rejection rate of non-target speech

Speaker Number	Number of passes	Number of failures	Rejection rate	Speaker Number	Number of passes	Number of failures	Rejection rate
1	0	10	1.0	8	1	9	0.9
2	0	10	1.0	9	0	10	1.0
3	1	9	0.9	10	2	8	0.8
4	1	9	0.9	11	0	10	1.0
5	2	8	0.8	12	1	9	0.9
6	0	10	1.0	13	0	10	1.0
7	0	10	1.0	Total	8	122	0.94

teaching. Excluding other noise by recognizing the target person’s vocal pattern is a feasible way to improve the sound quality of vocal teaching in online lessons. In this study, TDNN is improved using generative feature vectors to obtain a vocal pattern recognition system based on GTNN. The experimental results show that the feature vector extraction of GTDNN is better than TDNN and i-vector methods. The EERs of GTDNN are 11.3%, 12.0%, 4.9%, 6.2% and 6.1% in TIMIT speech bank, TIMIT speech bank with noise, Aishell speech bank, Aishell speech bank with noise and VoxCeleb speech bank, respectively, which are all in the lower level. the EERs of GTDNN for both short-time speech and long-time speech EER is the smallest among all four models, with the largest EER value of 9.6% and the smallest of 2.3%. The recognition pass rate for most of them is above 90%, and the overall recognition pass rate can reach 94%. the rejection rate of GTDNN is relatively high, with an overall rejection rate of 94% and a rejection rate of 90% for most impersonators. It can be seen that GTDNN has more stable recognition performance in various environments and can recognize the target speaker’s voice and exclude the interference of other non-target speaker’s voice, providing good vocal sound quality for online vocal teaching. Although GTDNN can extract more recognizable speaker voice features, its limitation is that it requires relatively high data volume, long training time for the model, and high equipment requirements. In the follow-up study, further optimization is needed for the learning ability of GTDNN. Future research could concentrate on improving the TDNN algorithms’ efficiency and accuracy. Experimenting with different neural network architectures, tuning hyperparameters, or using more advanced machine learning techniques to improve the system’s ability to handle diverse vocal patterns and accents could be part of this.

**Fundings.** The research is supported by: Hunan Philosophy and Social Science Foundation Project "Research on Inheritance and innovation of Suining traditional music culture under the background of integrated development of culture and tourism", Project No.: 18WTC28.

## REFERENCES

- [1] Bonfim, B., Bratfich, R. & Silva, M. HG Silva. *PREVISÃO DE CONSUMO DE ENERGIA UTILIZANDO REDE NEURAL COM RETARDO DE TEMPO (TDNN)*. *Colloquium Exactarum*. **12**, 63-70 (2021)
- [2] Cheng, G., Zhang, P. & Ji, X. Automatic Speech Recognition System with Output-Gate Projected Gated Recurrent Unit. *IEICE Transactions On Information And Systems E*. **102** pp. 355-363 (2019)
- [3] Fu, L. Discussion on the Differences between Theory and Practice in Vocal Music Teaching. *region - Educational Research and Reviews*. (2021)
- [4] Fu, L. Discussion on the Differences between Theory and Practice in Vocal Music Teaching. *Region - Educational Research And Reviews*. **3**, 6-9 (2021)
- [5] Fu, L. Research on the Reform and Innovation of Vocal Music Teaching in Colleges. *region - Educational Research and Reviews*. (2020)
- [6] Hanifa, R., Isa, K. & Mohamad, S. review on speaker recognition: technology and challenges. *Computers Electrical Engineering*. **90**, 1-10700 (2021)
- [7] Hanifa, R., Isa, K. & Mohamad, S. review on speaker recognition: technology and challenges. *Computers Electrical Engineering*. **90**, 1-10700 (2021)
- [8] He, Y. & Dong, X. Real time speech recognition algorithm on Embedded System Based on continuous Markov model. *microprocessors and Microsystems*. (0)
- [9] Hu, S., Xie, X., Cui M. J Deng, S. Liu S. Liu & S. Liu, J. Yu M Geng M Geng X Liu X Liu X Liu H Meng . *Neural Architecture Search For LF-MMI Trained Time Delay Neural Networks*. **30** pp. 1093-1107 (2022)
- [10] Kathania, H., Kadiri, S., Alku, P. & Others A formant modification method for improved ASR of children's speech. *Speech Communication*. **136**, 98-106 (2022)
- [11] Kethireddy, R., Kadiri, S. & Gangashetty, S. Deep neural architectures for dialect classification with single frequency filtering and zero-time windowing feature representations. *The Journal Of The Acoustical Society Of America*. **151**, 1077-1092 (2022)
- [12] Kuda, T. Aminu. Don , Kuda McGlinchey , Andrew Andrew Cowell Cowell . *Acoustic Signal Processing With Robust Machine Learning Algorithm For Improved Monitoring Of Particulate Solid Materials In A Gas Flowline*. *Flow Measurement Instrumentation*. **65** pp. 33-44 (2019)
- [13] Kumar, C. Hybrid models for intraday stock price forecasting based on artificial neural networks and metaheuristic algorithms. *pattern Recognition Letters*. (2021)
- [14] Kuo, T., Lo, R. & Chen, L. Shang-Hsien Cai Shang-Hsien Cai . *Activity Command Encoding Of Cerebral Cortex M*. **32**, 1-20500 (2020)
- [15] Smith, G., Heever, D. & Swart, W. The Reconstruction of a 12-Lead Electrocardiogram from a Reduced Lead Set Using a Focus Time-Delay Neural Network. *Acta Cardiologica Sinica*. **37**, 47-57 (2021)
- [16] Sun, J. Research on resource allocation of vocal music teaching system based on mobile edge computing. *computer Communications*. (2020)
- [17] Sun, J. Research on resource allocation of vocal music teaching system based on mobile edge computing. *Computer Communications*. **160**, 342-350 (2020)
- [18] Tatli, A., Kahvecioglu, S. & Karakoc, H. Time-Series Prediction for Amount of Airworthiness Based on Time-Delay Neural Networks. *Elektronika Ir Elektrotechnika*. **2020**, 28-32 (0)
- [19] Tomlinson, J. Music therapist collaboration with teaching assistants for facilitating verbal and vocal development in young children with special British Journal of Music Therapy. (2020)
- [20] Yang, H. Quality Analysis of Multimedia Teaching in Vocal Music Class Combining Elitist Teaching-Learning-Based Optimization Algorithm. *Hindawi*. **4781**, 1-65947 (2021)
- [21] Zekhnine, C. & Berrached, N. Human-Robots Interaction by Facial Expression Recognition. *international Journal of Engineering Research in Africa*. 2020. (0)
- [22] Zhu, W., Jin, H., Chen J. L Luo, J. Wang J. Wang, J. Wang Q. Lu & Lu, Q. Lu A Li . *A Hybrid Acoustic Model Based On PDP Coding For Resolving Articulation Differences In Low-resource Speech Recognition*. **192** pp. 4 (2022)
- [23] Lee, S., Abdullah, A. & Jhanjhi, N. A review on honeypot-based botnet detection models for smart factory. *International Journal Of Advanced Computer Science And Applications*. **11** (2020)
- [24] Azeem, M., Ullah, A., Ashraf, H., Jhanjhi, N., Humayun, M., Aljahdali, S. & Tabbakh, T. Fog-oriented secure and lightweight data aggregation in iomt. *IEEE Access*. **9** pp. 111072-111082 (2021)
- [25] Gaur, L., Solanki, A., Wamba, S. & Jhanjhi, N. *Advanced AI techniques and applications in bioinformatics*. (CRC Press,2021)

*Edited by:* Kumar Abhishek

*Special issue on:* Machine Learning and Blockchain based Solution for Privacy and Access Control in IoT

*Received:* Sep 28, 2023

*Accepted:* Jan 14, 2024



## APPLICATION OF GENETIC ALGORITHM IN OPTIMIZATION SIMULATION OF INDUSTRIAL WASTE LAND REUSE

PENG BAI\*, YUNAN ZHAO<sup>†</sup> AND JUNJIA CHANG<sup>‡</sup>

**Abstract.** In order to better understand the application of optimization simulation for industrial waste land reuse, the author proposes an application study based on nonlinear genetic algorithm in the optimization simulation of industrial waste land reuse. The author takes the landscape renovation and reuse of industrial waste sites as the research object, and through research on the current situation of landscape renovation and reuse of industrial waste sites both domestically and internationally, as well as on-site inspections, attempts to use landscape design techniques to deal with this once glorious but destructive industrial landscape that has already declined. Secondly, a genetic algorithm for enhancing the timeliness of industrial waste land reuse is proposed, which is based on random walks, combine users' long-term and short-term preferences to calculate the most suitable Top-N industrial waste land reuse optimization model for the current period. Finally, the two algorithms proposed by the author were experimentally validated on the dataset. In the CiteU Like dataset, the best performance was achieved at  $a=0.4$ , while in the JD dataset, the best performance was achieved at  $a=0.6$ . When  $k=6$ , the hit rate significantly decreases by about 50%. The URT-R genetic algorithm exhibits a high recommendation hit rate in recommendations targeting timeliness. The author analyzed the different characteristics of industrial waste reuse in scenic areas and optimized their essence and transformation methods, further improving the transformation and renewal methods of industrial waste land in the process of urban development in China. I hope to provide useful references for future research on related topics and practices.

**Key words:** Genetic algorithm, reuse of industrial waste land, optimization

**1. Introduction.** Waste industry refers to the land used for industrial production and transportation, transportation, storage, and then abandonment, such as landfills, quarries, factories, railway stations, loading stations, industrial wastes, etc. The wasteland industry has increased with the decline of the traditional industry, which is caused by the impact of the industrial disaster on the environment, as well as the development of modern technology, the transformation of industry, and the transfer of resources. Therefore, people are increasingly aware that although industrial civilization has brought certain material wealth to humanity, it has also brought a series of environmental problems: Environmental pollution, land desolation, species loss, Balance of nature imbalance, and a series of social problems: waste of land resources, serious impact on the living standards of residents around industrial wasteland, and so on. The emergence of these problems has sparked social attention to industrial waste land. How to achieve ecological restoration and seek scientific solutions in heavily polluted waste land caused by human overuse has become a practical topic in the research of industrial waste land regeneration, and has attracted attention from all sides. The reuse of industrial waste land has become a global topic of common concern.

In the 1960s, more and more landscape designers and ecologists began to pay attention to the reuse of industrial waste land, due to the reasons of the times, this focus is no longer satisfied with the simple focus on land reuse in the past, but his analysis has already shifted to conforming to historical and cultural values as well as the spiritual nature of this abandoned land. American city historian Lewis Mumford once said, "If you want to make a better living in a new city, we must first understand the historical characteristics of this city. People gradually realize that the successful reuse of industrial waste land from the historical value of buildings and derived cultural emotions is also a way to maintain people's experience of the historical characteristics of cities. Therefore, in the process of continuous transformation and utilization of industrial waste land, there has been a bold attempt to use landscape methods and approaches to solve the problem of reuse. After the

---

\*Shenyang Normal University, College of Fine Arts and Design, Shenyang, Liaoning 110034, China ([PengBai939@163.com](mailto:PengBai939@163.com))

<sup>†</sup>Shenyang Normal University, College of Fine Arts and Design, Shenyang, Liaoning 110034, China (Corresponding author, [YunanZhao7@163.com](mailto:YunanZhao7@163.com))

<sup>‡</sup>Dalian University of Foreign Languages, School of Software, Dalian, Liaoning 116044, China ([JunjiaChang@126.com](mailto:JunjiaChang@126.com))

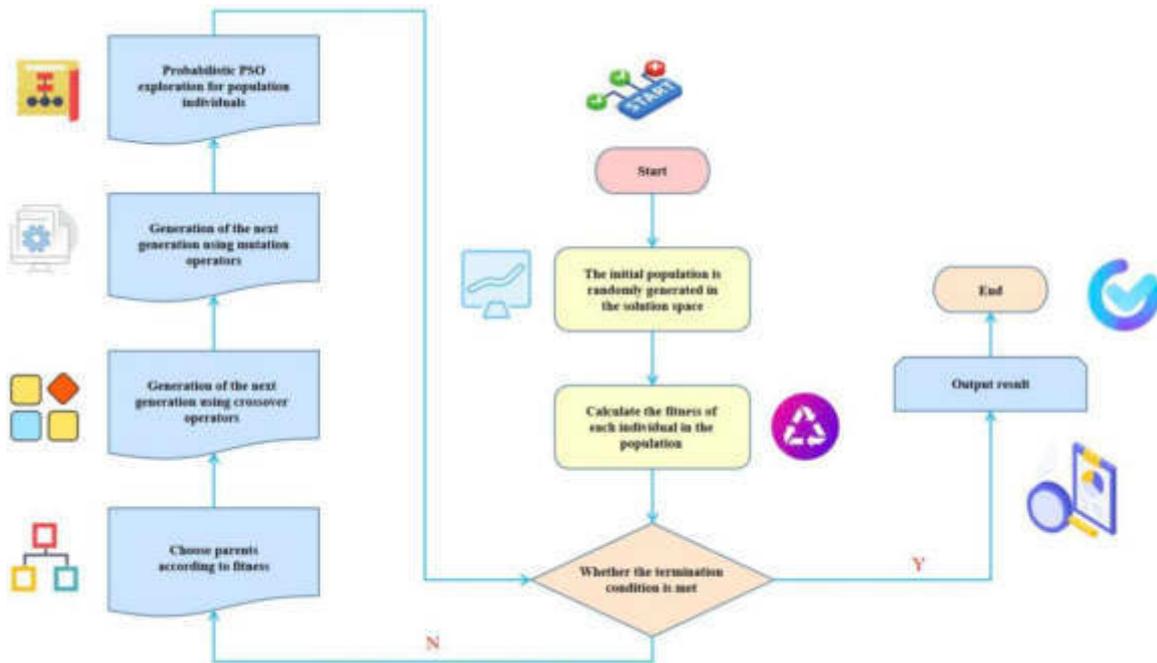


Fig. 1.1: Optimization simulation of industrial waste land reuse

1970s, with the decline of the traditional economy, the strengthening of environmental awareness, and the increase of environmental mobility, the number of industrial upgrades gradually increased. The continuous development of science and technology, as well as the achievements of ecology and biotechnology, have also provided technical support for the transformation of industrial wastes. In 1972, the Seattle Natural Gas Production Plant in the United States was the precondition for reusing the industrial wastes using the landscape design method. It has had a general impact on the landscape design in terms of park form, industrial landscape style, and cultural value. In the 1990s, people tried to use landscape design techniques to solve the historical puzzlement of the landscape industry, which has caused the ecological environment deterioration and has fallen behind, resulting in a large number of design patterns [15, 14]. Designers have utilized a comprehensive approach of science and art to achieve the goals of environmental renewal, ecological restoration, cultural reconstruction, and economic development in industrial wasteland. As shown in Figure 1.1:

**2. Literature Review.** Industrial waste land, in a broad sense, was once used for industrial production, with the continuous growth of technology and urban renewal, many sites have been forced to transfer to the surrounding areas of the city, resulting in many industrial waste sites, most of these industrial legacy sites have serious industrial pollution, and the industrial traces are very obvious. It includes many types, including abandoned storage equipment, industrial equipment, and open spaces formed by long-term storage of materials containing pollutants.

Most of this part of the land no longer has useful value, and the heavily polluted part is difficult to reuse, resulting in a series of ecological and economic problems. Industrial wasteland is a part of Urban culture economy, which has the function of extending the urban context and belongs to Urban culture heritage. Although the emergence of industrial waste land appears to be an economic recession, it has generated a series of social and environmental problems. In order to find out the treatment methods of different industrial wastelands is a problem that we have been exploring. It should not only conform to the local cultural background and be able to be well integrated into the urban development to conform to the urban development, but also play a positive role in the environment. It can improve the air quality of the park, reduce noise, and

beautify the environment. Economically, reducing the waste of land resources has established a positive urban image and promoted economic development, bringing economic benefits through resource development. In society, the impact of the environment on urban development has been reduced, people's living space has been optimized, and people's needs for urban environmental development have been met. Leirpoll, M. E. proposed that ecological reuse is a rigorous test of ecological theories, focusing on the characteristics, destruction processes, and restoration methods of ecosystems [8]. Hu, Z. B. believes that ecological reuse refers to a series of management activities that use modern restoration methods to restore damaged ecosystems to their pre damaged state [6].

### 3. Methods.

**3.1. Design principles and methods for the reuse of industrial waste land.** Respecting the natural growth state and using the above ecological and ecological restoration methods to transform the abandoned land, the first thing to do is to preserve the characteristic Natural landscape of the site. Continuing to update the visual landscape in the venue for people to enjoy and rest, the combination of functionality and visual landscape has jointly created a new way of life [11, 16]. Based on the above analysis and summary, the reconstruction and restoration of industrial waste land can be adjusted from the following five principles: Ecological priority principle, respect for the site principle, adaptation to local conditions principle, people-oriented principle, and harmonious coexistence principle.

#### (1) Ecological priority principle

Firstly, summarize the specific reasons for the polluted environment on the site and carry out targeted adjustments and reconstruction, it is best to give the site a brand new look with minimal mobilization to ensure the balance of the natural ecosystem [12]. Secondly, on the basis of ensuring the continuity of the entire ecology, the maximum effort should be made not to cause secondary damage, the principle of fully utilizing the original terrain and landforms of the site to not damage the existing good vegetation style of the site is to use leftover materials and usable materials for design, reconstruction, and restoration to achieve sustainable development of the site. The Shanghai World Expo Houtan Wetland Park is a clear case in point. There are approximately 16 hectares of wetlands in the park, which can accommodate many wild birds as their habitats, the original old park's wetlands are mostly isolated with hard paving around them, the rebuilt park has abandoned the hard paving and replaced it with wild plants suitable for wetland growth, which has prevented wetland flooding and greatly improved the ecological integrity of the entire park [4].

#### (2) Respect the principle of venue

The regeneration of abandoned land should first consider respecting the characteristics of the site, in order to maintain sustainable ecological stability, it is necessary to start from the natural laws of plant growth and protect the plant communities left over from the original site, and try not to damage its existing ecological environment. The preservation of distinctive landscapes and the integration of fresh cultural enterprises enable the long-term development of the park, and drive the economic benefits of the entire park. For example, the Qijiang Park in Zhongshan City, formerly a waste shipyard in Zhongshan City, is also polluted to a corresponding degree in different scales, adhering to the principle of restoring the ecological space of Urban green space and inheriting industrial history and culture [9]. The park retains representative landmarks from the original shipyard as symbols of the new park, including the preservation of plants and structures such as railway tracks in the original factory area.

#### (3) Principle of adapting to local conditions

Adapting to local conditions in a broad sense means renovating the landscape to varying degrees based on the local natural and cultural characteristics, including different areas, the diversity of surrounding users and the type of venue to be updated. This requires effective analysis based on the actual local situation during design, and different methods should be adopted to update and rebuild the landscape [3, 1].

Firstly, before updating and designing, it is necessary to conduct a firsthand investigation and analysis of the surrounding environment, the economic situation of the surrounding users, and the current situation of surrounding pollution. Based on the pollution situation in the park and the analysis of the problems and advantages in the surrounding terrain, an effective method is proposed. Secondly, after meeting the needs of different locations and regions, it is more important to consider the ecological status of the park in the next

ten to twenty years, different plots or industrial waste land have diverse characteristics, and no factory area is copying the achievements of others.

For example, if the pollution level in the factory area is relatively low and the geographical location of the park is relatively superior, the buildings that are preserved are relatively intact. Based on the actual situation, they can be transformed into projects that can drive economic development, such as art museums, historical museums, and other enterprises of different scales. Not only can it promote the development of the surrounding economy, but it also makes reasonable use of various factors available in the factory area. As an important design principle, adapting to local conditions stipulates that designers must be people-oriented and start from their own unique perspective, aiming to establish urban public green spaces that are more suitable for a region [10, 17].

**3.2. Design method for industrial waste land transformation and reuse.** The transformation of industrial waste land into urban public green spaces requires considering many factors that are relatively difficult to design, when designing for renovation and reuse, attention should be paid to the following aspects:

(1) The preservation and renovation of abandoned industrial buildings, shopping and industrial facilities, most of the remaining machinery or structures in the factory area that have original industrial park functions, how to effectively preserve and reuse them is a key part that we should pay attention to when designing, achieving a unified style between preservation and reconstruction [2].

One is overall preservation. The preservation of this part mostly refers to buildings with important usage functions in the park or well preserved buildings and structures in the park. After the renovation, the overall production process of the original industrial park can be further felt.

The second is selective retention. When encountering structures that can be preserved or have preservation value, appropriate and selective preservation can be carried out based on the principle of preserving the style and features of industrial sites. These parts are mostly industrial buildings with obvious characteristics or well preserved buildings [7].

The third is the preservation of components. There are a large number of structures left behind in the park that have been repaired and can be reused. This not only updates the landscape of the factory area, but also allows people to truly feel the original industrial style and the sense of the times brought by history.

(2) Treatment of surface relics after industrial production

Regardless of the type of industrial production, traces will be left on the surface to some extent, however, there are many unique industrial surface relics that can be effectively protected to form a unique landscape. For ordinary surface relics, preservation and restoration methods can be chosen to enhance their overall industrial style as a new landscape.

(3) Waste utilization and pollution treatment

The waste and unused raw materials on the site can be reused, provided that they do not cause further pollution to the park, for materials that pollute the environment, they can be recycled through technical treatment methods. Ultimately, choosing local materials according to local conditions reduces economic pressure while also increasing ecological benefits.

(4) Soil testing

Conduct professional testing on the soil of different plots in the park, and develop different remediation methods based on specific circumstances. It is recommended to transfer the most severely polluted soil. By analyzing the current situation of soil and using technical techniques to identify the main sources and pollutants of pollution, targeted remediation can be carried out. From the local to the overall, gradually forming a harmonious ecosystem, using different treatment methods to achieve good soil planting status.

**3.3. Genetic algorithm recommendation method for reuse of abandoned land in industrial scenic spots.** The author proposes a genetic algorithm for the reuse of abandoned land in industrial scenic areas. The algorithm mainly consists of two steps: a. Combining user stable preferences with short-term preferences generated by external stimuli to calculate the user's preference for scenic spots; b. Design a ranking method for recommended scenic spots, recommending the top N favorite scenic spots to users. Next, we will explain step (1) of the algorithm through the definition of the UPT graph, and then introduce the sorting method.

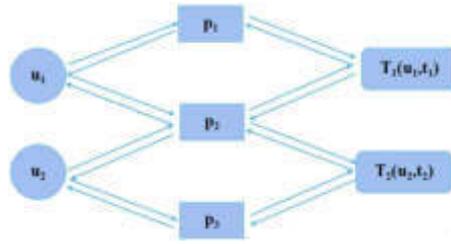


Fig. 3.1: Example of UPT diagram

Definition 2: (User Scenic Area Time Period Relationship Graph, UPT Graph) A UPT graph is a directed tripartite graph  $G (U, P, T, E, w$ , where:  $U$  is the set of user  $u$ ;  $P$  is the scenic area  $P$ ; Set of;  $T$  is the set of time periods  $T (u, t)$ .  $E$  is the set of edges in the graph, with edges between  $u$  and  $p$ , if and only if user  $u$  has visited  $p$ ,  $p$ ; There is an edge between and  $T$ , if and only if user  $u$  has visited the scenic area  $p$  during time  $t$ ;  $W$  represents the weight of the edge, which is defined as follows formula (3.1):

$$w < v_i, v_j \geq \begin{cases} \frac{N(p_i)}{\sum_{i=1}^n N(p_i)} \\ \frac{1}{\sum_{i=1}^n N(u_i)} \\ w_p \\ w_t \end{cases} \tag{3.1}$$

A UPT graph is a directed graph, with weights  $w < v_i$  and  $v_j >$  calculated separately based on the vertices associated with the directed edge. The weight value of edge  $< u_j, p_i >$  measures the long-term preference of user  $u_i$  for scenic spot  $p_i$ , which is calculated from the proportion of  $p_i$  scenic spots in user  $u_i$ 's historical data; The weight of edge  $< u_i, p_i >$  measures the special degree of the scenic area  $p_i$  to the user  $u_i$ , calculated by the reciprocal of the number of visitors to the scenic area  $p_i$ . The higher the value, the more special the  $p_i$  is to the  $u_i$ . Similarly, during a certain period of time, the user's preference for the scenic area and the special degree of the scenic area to the user are expressed in  $w_p$  and  $w_t$  respectively.

Figure 3.1 shows the UPT diagram composed of users, scenic spots, and time period nodes. The similarity between scenic spots can also be calculated through this graph. If two scenic spots are associated through the user node  $U_I$ , their similarity can be calculated based on the user's long-term preferences; If two scenic spots are associated through the time period node  $T_i$ , their similarity can be calculated based on users' short-term preferences.

After the construction of the UPT graph is completed, a new recommendation method is designed based on the idea of random walk, taking into account the characteristics of nodes in the UPT graph [5]. The main idea of the R-UPT method is to start from the node  $u_i$  and follow the edges between the vertices, then randomly walk to an unknown scenic spot, which matches the user's long-term preferences; On the contrary, if you randomly walk from node  $T_i$ , you can reach a scenic spot that matches the user's short-term preferences. The user's preference for a certain scenic spot  $p_i$  can be calculated by taking the  $u_i$  node or  $T_i$  node as the tail of the arc and the  $p_i$  node as the end of the arc, the scenic spot  $p_i$  with the maximum sum of edge weights will be recommended to the user  $u_i$ . Given that a user  $u_i$  starts from a user node or a time period node, there are many paths to reach the scenic spot node  $p_i$ , the algorithm only considers the shortest path to reach the scenic spot node from the user node or time period node, which is a path with a length of 3, there are four situations where the shortest path to reach the scenic spot node on the UPT diagram is as follows:

a. User  $\rightarrow$  Scenic Area  $\rightarrow$  User  $\rightarrow$  Scenic Area: Starting from user node  $u_1$ , reach a certain scenic area  $p_2$  that the user has browsed, and then reach scenic area  $p_3$  through user  $u_2$  who has browsed the scenic area, as shown in the path  $u_1 \rightarrow p_2 \rightarrow u_2 \rightarrow p_3$  in the Figure;

b. User  $\rightarrow$  Scenic Area One Time Period  $\rightarrow$  Scenic Area: Starting from user node  $u_1$ , reach a certain scenic area  $p_1$  that the user has browsed, and then pass through time period node  $T_1$  to reach scenic area node  $p_1$ , as shown in the path  $u_1 \rightarrow p_1 \text{ one} \rightarrow T_1 \text{ one} \rightarrow P_2$  in the Figure;

c. Time period  $\rightarrow$  scenic area  $\rightarrow$  user  $\rightarrow$  scenic area: Starting from time period node  $T_1$ , reach scenic area  $p_1$ , and reach scenic area  $p_2$  through user  $u_1$  who has browsed scenic area  $p_2$ , as shown in the path  $T_1 \rightarrow p_1 \rightarrow u_1 \rightarrow P_2$  in the Figure;

The user's preference level  $r(u_i; p_i;)$  for a certain scenic area can be expressed as equation (3.2):

$$r(u_i, p_i) = \sum w < v_i, v_j > \quad (3.2)$$

Using PageRank's improved algorithm to calculate the ranking of scenic spots, we recommend Top  $n$  users' favorite scenic spots. Formula (3.3) shows the calculation method for sorting nodes in the graph.

$$PR = \alpha \cdot M \cdot PR + (1 - \alpha) \cdot \tilde{d} \quad (3.3)$$

In order to achieve the idea of recommending Top  $N$  scenic spots to each user, a personalized vector is first constructed for each user, and then the scenic spot nodes in the UPT graph are sorted according to formula (3.3), the Top- $N$  node is recommended to users, the algorithm is based on Breadth-first search and implemented by queue structure, after walking to a node, the sum of path length and path weight is updated; Finally, call the sorting function to find Top- $N$  scenic spots to be recommended.

**4. Experimental Analysis.** The author's experiment was conducted on two datasets, CiteULike and JD. This experiment divides the dataset into a test set and a training set, the user's most recent purchase record is used as the test set CUL-Test, JD Test, and the remaining purchase records are used as the training set CUL-Train, JD Train. The experimental environment is Windows 7 operating system, Intel i5 3.3G CPU, and 8GB memory.

Measurement parameters:

The author uses Hit Ratio (HT) to measure the accuracy of Top  $N$  recommendations. Firstly, generate a top  $N$  recommended product list for user  $u$  during the time period  $t$ , denoted as  $P(u, t)$ . If a product appears in the top  $N$  recommended product, it is recorded as a hit [13]. The hit rate formula (4.1) measures:

$$HT = \frac{\sum_u (p_u \in P(u, t))}{|U|} \quad (4.1)$$

From Figure 4.1(a), it can be seen that with the  $\beta$  as the value changes, the hit rate of algorithm URT-R increases first and then decreases, stay  $\beta$  when the value is set to 0.6, it approaches the peak. Due to the relationship between the other three algorithms and  $\beta$  it is irrelevant, therefore, the data will not fluctuate, as can be seen, without considering timeliness, the hit rate of the CF-T algorithm is between UserKNN and ItemKNN.

Figure 4.1(b) shows that on the JD dataset, the hit rate varies with  $\beta$  the overall trend of changes in values is similar to Figure 4.1(a), only corresponding to the peak value  $\beta$  the values are different. Due to the high density of the JD dataset, its average hit rate is higher than that of the CiteULike dataset.

In addition, the author also focuses on the coordination parameters  $\alpha$  conducted experiments, stay  $\beta=0.5$ . Under the premise of 0.5, change  $k$  and  $\alpha$  observe the change in hit rate based on the value of, and the experimental results are shown in Figure 4.1(b).

This article will compare the genetic algorithm proposed by the author with two classic algorithms: the user-based filtering algorithm UserKNN and the object-based collaborative filtering algorithm ItemKNN. and the object-based collaborative filtering algorithm ItemKNN. Firstly, in  $\beta$  Testing the changes of the attribute value while taking different values and processing the algorithm of the two datasets. The experimental results are shown in Figure 4.2(a)(b).

The experiment in Figure 4.2 shows the coordination parameters  $\alpha$  the relationship with hit rate is best in the CiteU Like dataset at  $\alpha=0.4$ , while in the JD dataset at  $\alpha=0.6$ . When  $k=6$ , the hit rate significantly decreases by about 50%.

**5. Conclusion.** In response to the strong timeliness of optimization simulation for the reuse of abandoned industrial scenic spots, the author proposes two genetic algorithms, CFT and URT-R. The CF-T algorithm improves the traditional system filtering algorithm and the definition of similarity between users and industrial

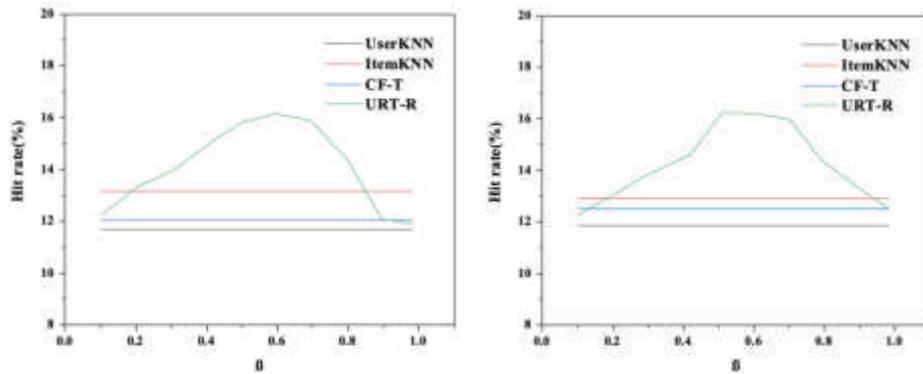


Fig. 4.1: (a) Hit rate  $\beta$  Changes in values (Cite ULike dataset); (b) Hit rate  $\beta$  Changes in values (JD dataset)

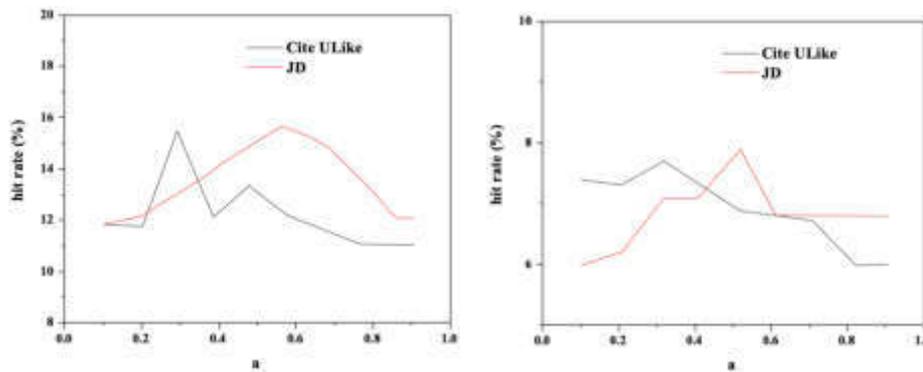


Fig. 4.2: (a) Coordination parameters  $\alpha$  Relationship with hit rate ( $k=10$ ); (b) Coordination parameters  $\alpha$  Relationship with hit rate ( $k=5$ )

scenic spots, and adds time factors to obtain the recommended Top N industrial scenic spots; The URT-R algorithm is based on random walks, combining users' long-term and short-term preferences, calculate the Top-N algorithm that is most suitable for users at the current time, and recommend it through optimization simulation of industrial waste land reuse. Finally, experimental verification was conducted on the dataset for the two algorithms proposed in this article. In terms of timeliness recommendations, the URT-R genetic algorithm showed a high recommendation hit rate, further improving the recommendation accuracy on the data.

**Acknowledgement.** This work was supported by Liaoning Social Science Planning Fund Project (L21BGL052).

#### REFERENCES

- [1] S. R. ABELLA AND T. A. SCHETTER, *Variation in characteristics and conservation values of plant communities on abandoned agricultural lands with and without fires*, Applied Vegetation Science, 24 (2021), p. e12629.
- [2] A. BEN, E. PUSHPALAKSMI, J. SAMRAJ, AND A. GURUSAMY, *Biodegradation of tannery effluent dyes by using aspergillus niger isolated from the tannery effluent and reuse of biotreated water in agricultural field*, Journal of Biodiversity and Environmental Sciences, 18 (2021), pp. 74–80.
- [3] D. CHEN, B. ZHANG, C. FADDA, D. JARVIS, N. BERGAMINI, G. HAN, M. ZHAO, K. BAI, AND Z. ZHANG, *Spontaneous grassland recovery on abandoned croplands in northern china: Different vegetation patterns in desert and typical steppe*, Science of The Total Environment, 790 (2021), p. 148155.
- [4] P. COSTA, E. BRITO-HENRIQUES, AND C. CAVACO, *Interim reuse in urban derelicts: Uncovering the community's attitudes and preferences through scenario-elicitation*, Cities, 111 (2021), p. 103103.

- [5] M. DI NARDO AND T. MURINO, *The system dynamics in the human reliability analysis through cognitive reliability and error analysis method: A case study of an lpg company*, Int. Rev. Civ. Eng, 12 (2021), pp. 56–68.
- [6] A. HU, S.-H. CHANG, X.-J. CHEN, F.-J. HOU, AND Z.-B. NAN, *Temporal heterogeneity has no effect on the direction of succession in abandoned croplands in a semiarid area of northwest china*, Land Degradation & Development, 32 (2021), pp. 91–100.
- [7] A. KALELI AND H. İ. AKOLAŞ, *The design and development of a diesel engine electromechanical egr cooling system based on machine learning-genetic algorithm prediction models to reduce emission and fuel consumption*, Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science, 236 (2022), pp. 1888–1902.
- [8] M. E. LEIRPOLL, J. S. NÆSS, O. CAVALETT, M. DORBER, X. HU, AND F. CHERUBINI, *Optimal combination of bioenergy and solar photovoltaic for renewable energy production on abandoned cropland*, Renewable Energy, 168 (2021), pp. 45–56.
- [9] D. LI, X. ZHANG, J. A. DUNGAIT, X. WEN, T. A. QUINE, AND Q. WANG, *Changes in the biological n<sub>2</sub>-fixation rates and diazotrophic community as vegetation recovers on abandoned farmland in a karst region of china*, Applied Soil Ecology, 158 (2021), p. 103808.
- [10] ———, *Changes in the biological n<sub>2</sub>-fixation rates and diazotrophic community as vegetation recovers on abandoned farmland in a karst region of china*, Applied Soil Ecology, 158 (2021), p. 103808.
- [11] J. S. NÆSS, C. M. IORDAN, H. MURI, AND F. CHERUBINI, *Energy potentials and water requirements from perennial grasses on abandoned land in the former soviet union*, Environmental Research Letters, 17 (2022), p. 045017.
- [12] A. V. PRISHCHEPOV, E. V. PONKINA, Z. SUN, M. BAVOROVA, AND O. A. YEKIMOVSKAJA, *Revealing the intentions of farmers to recultivate abandoned farmland: A case study of the buryat republic in russia*, Land Use Policy, 107 (2021), p. 105513.
- [13] M. QIN, *Evolution of labor relations in the development of human resources based on improved genetic algorithm*, Journal of Circuits, Systems and Computers, 31 (2022), p. 2250272.
- [14] G. RONG, H. WU, P. YANG, G. DUAN, X. SHEN, N. GE, AND X. WEI, *Dynamics of new-and old-soil organic carbon and nitrogen following afforestation of abandoned cropland along soil clay gradient*, Agriculture, Ecosystems & Environment, 319 (2021), p. 107505.
- [15] L. WANG, Q. TIAN, C. WU, H. HE, C. GUO, J. WANG, AND H. ZHAO, *Study on soil microorganism of abandoned land in qingtu lake in different years*, Journal of Environmental Protection and Ecology, 22 (2021), pp. 1408–1415.
- [16] R. WANG, J. ZHANG, H. SUN, S. SUN, G. QIN, AND Y. SONG, *Effect of different vegetation on copper accumulation of copper-mine abandoned land in tongling, china*, Journal of Environmental Management, 286 (2021), p. 112227.
- [17] G. YANG, R. XU, Y. CHEN, Z. WU, Y. DU, S. LIU, Z. QU, K. GUO, C. PENG, J. CHANG, ET AL., *Identifying the greenhouses by google earth engine to promote the reuse of fragmented land in urban fringe*, Sustainable Cities and Society, 67 (2021), p. 102743.

*Edited by:* B. Nagaraj M.E.

*Special issue on:* Deep Learning-Based Advanced Research Trends in Scalable Computing

*Received:* Aug 3, 2023

*Accepted:* Oct 17, 2023



## SPACE LAYOUT SIMULATION OF ASSEMBLED NANOARCHITECTURE BASED ON IMPROVED PARTICLE SWARM OPTIMIZATION

HUAN HUANG\*

**Abstract.** In order to solve the problem that the traditional building space configuration model cannot meet the optimization of building space characteristics, the author proposes the optimization of building space utilization based on spatialized particle swarm optimization. First, to solve the problem of optimal allocation of space units, the PSO algorithm is modified to encode the space units by means of character coding. Secondly, the maximum standardization method is used for data processing, and the factors affecting space utilization are summarized, the objective function of optimal allocation of architectural space is given from three aspects: economic benefits, social benefits and ecological benefits; Finally, by analyzing the advantages and disadvantages of master-slave parallel model and point-to-point parallel model, a chained parallel structure is proposed. The experimental results show that: The experimental data is based on the utilization of building space in these three regions in 2015, and the vector map is divided into  $30\text{ m} \times A$  grid of  $30\text{ m}$  in size, and all statistical data and spatial data are projected on each grid cell. The difference between the fitness values of the final convergence of the three parallel models is small, and the main difference is the convergence speed. During the run time test, set the three parallel models to run under the conditions of 8, 16, 32 and 64 nodes respectively. Because of the combination of the advantages of master-slave model and point-to-point model, the running time of chained parallel model is significantly lower than that of the other two parallel models. Conclusion: Through data simulation test, it is verified that the chained parallel model has higher fitness, convergence speed and shorter running time, and its performance is better than the other two, indicating that the optimization algorithm proposed by the author has good performance.

**Key words:** Spatialized particle swarm, Building space utilization rate, Symbol coding, Maximum standardization method, Chain parallel model, Master-slave parallel model, Point to point parallel model

**1. Introduction.** Land use planning is based on the natural and historical characteristics of regional land and social and economic development, comprehensive technical and economic measures to plan and arrange the future land use of the region in advance, reasonably allocate land resources in time and space, and reasonably organize land use. In order to promote sustainable land use, the state has established a national, provincial, municipal, county and township land use planning system. The five level planning model gradually controls the land use structure from top to bottom, and organizes from macro to micro, promotes land use. The main and difficult problem in land use planning is the optimal utilization of land, including the optimization of land use and structure, as well as the optimization of land use, it is not only necessary to determine the proportion of land used by various departments, but also to implement the specific spatial location. Through the optimization of land allocation, determine the best land use ratio, quantity and location of land, that is, solve the problems of “what”, “how much”, “where” in land use planning. The optimal allocation of land use is also a multiple optimization problem. According to the difference of region, the optimization goals are win-win and conflict. In the process of optimization, it is necessary to maintain the interests of all parties, balance the land resources and needs, and maximize the social, economic, and ecological benefits of the regional land use.

The optimal allocation of land use is a hot issue in land use planning. Optimization of the multi-scale structure of land use has achieved the selection of optimization goals and the application of optimization methods, but the research on the theory and methods of optimization of land use in general is a little weak. Based on intelligent optimization algorithms, many scholars have proposed various land use optimization allocation models. However, these models have various problems, which make them stay in the stage of scientific research and fail to be applied to the practice of land use planning. These problems include: The degree of spatialization of the model is not enough, and the spatialization operation is inconvenient; The integrated domain knowledge is weak, which affects the optimization effect of the model; The operation efficiency is low, and it is difficult to apply to high-precision land use optimal allocation in large areas. Study how to make intelligent optimization

---

\*Wuhan Institute of Shipbuilding Technology, Hubei, wuhan, 430070, China

algorithm spatialized, knowledge-based and parallelized, it will greatly improve the optimization effect and efficiency of the land use optimization model, and improve the practicality of the intelligent optimization model.

**2. Literature Review.** Housing is the basic demand of people all over the world, with the progress of science and technology and social development, the housing construction mode is also improving in exploration, inspired by the booming development of manufacturing industry, people begin to think about whether the construction mode can also take the road of industrialization. The construction industry has always been dominated by cast-in-situ construction [11]. In recent years, Beijing, Shenyang, Shenzhen, Shanghai, Guangzhou, Hong Kong and Taiwan have responded to the requirements of the modernization of the national construction industry, take the lead in promoting the development of prefabricated buildings. Ghahramani, M. believes that off-site manufacturing and standardized production can effectively promote the development of the construction industry in terms of quality, cost, progress, safety and energy efficiency, and has designed and described this to guide the prefabrication and assembly technology [4]. After analyzing the current situation of low efficiency in the construction industry, de Oliveira proposed that prefabrication can be adopted in the construction, based on this, he analyzed different construction projects after adopting the assembly method, and the data obtained proved that the assembly method can indeed reduce a lot of waste [3]. Song, M. pointed out that building prefabrication is the future trend of the industry, and concluded that prefabricated buildings can indeed bring more obvious economic and social benefits through economic, environmental, social and other assessments of prefabricated building projects [13]. Carri ó n, D. pointed out the importance of the layout of temporary materials for the layout management of the construction site, and proposed to quantify the use frequency of a material in the whole project period by the material accessibility level, that is, the daily use of a material, the three key factors affecting the use frequency are: Material handling, transportation route and transportation distance.

Through the data simulation test, the author has verified that the chained parallel model has higher fitness, convergence speed and shorter running time, and its performance is better than the other two, indicating that the optimization algorithm proposed by the author has good performance.

### 3. Research Methods.

**3.1. Spatialization of particle swarm optimization algorithm.** In view of the shortcomings of the existing building utilization configuration model in the degree of spatialization, the author adopts the spatialized coding method, operation unit and operation operator. The mapping from numerical space to geographical space is realized based on particle swarm optimization algorithm, and the optimal configuration model of building space utilization is constructed according to rules and multi-agent principle. Finally, the parallel improvement of the particle swarm optimization model is realized by the method of sub cluster division. Its structural framework is shown in Figure 3.1.

For the application of PSO to the optimal layout model of domestic space applications, it is necessary to verify the relation between the parameters of PSO and the parameters of the optimum layout model for the domestic space applications, furthermore, it is necessary to study the application of the PSO to the optimal configuration problem. In order to separate the responsibilities of algorithm engineers and building space planners, the corresponding relationship between the optimal configuration parameters of building space utilization and the parameters in particle swarm optimization should be as comprehensive and accurate as possible. Particle swarm optimization has four parameters: Particle, particle position, particle speed and particle fitness.

In particle swarm optimization, particles are used to represent candidate solutions of the problem to be solved. For the problem of optimal configuration of building space utilization, particles can be defined as the configuration scheme of building space utilization. The particle swarm consisting of all particles represents a collection of multiple configuration schemes, and the particle position is used to represent the value of the building space allocation unit and the use type of the building, generally, the array  $(x_1, x_2, \dots, x_n)$  is used to represent the position of particles, each element in the array represents a dimension, the variable  $x$  in the dimension corresponds to the numerical value of the building space allocation unit. Because of the specific definition of particles, their speed in each dimension is defined as the direction of building unit usage conversion corresponding to that dimension. In general, the transformation probability is used to describe the velocity

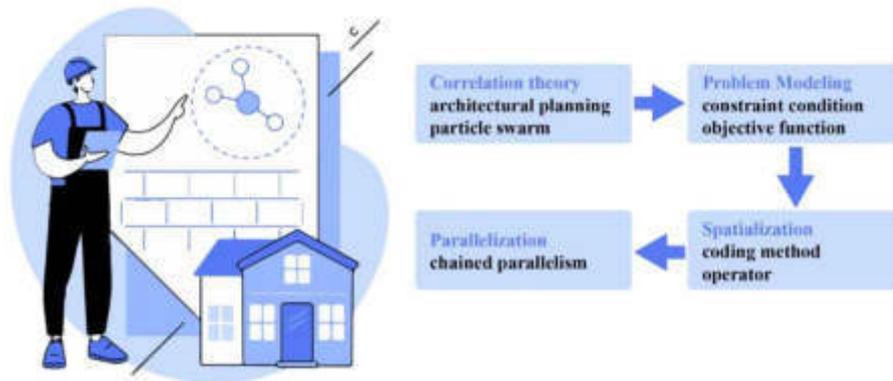


Fig. 3.1: Structure framework of spatialized particle swarm optimization algorithm

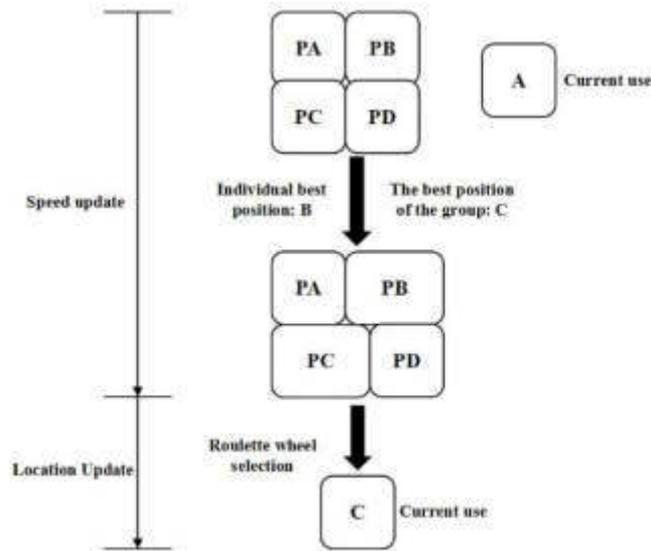


Fig. 3.2: Schematic diagram of particle velocity and position update based on symbol encoding

of particles. The fitness indicates the advantages and disadvantages of the optimized configuration scheme for the utilization of the building cluster, under the limited land and use space, the objective function should be constructed from multiple perspectives, then the fitness function is determined.

For the problem of building space configuration, it is necessary to integrate the characteristics of three-dimensional space for particle swarm optimization, mainly including the spatialization of coding methods, operation units and operators. The author uses symbol - based coding method to encode PSO spatially. Residential buildings are classified into three categories: residential, commercial, industrial, and public, and the values A, B, C, and D, respectively. In PSO, the location of the particle in length is A. The particle velocity is represented by the probability vector  $[PA, PB, PC, PD]$ , generally, the initial probability vector elements of each particle have the same value. The value change of the particle will be affected by the best historical position of the individual and the best historical position of the population, so the probability of each dimension will change. Finally, the purpose of the building space is finally determined through roulette selection, as shown in Figure 3.2.

### 3.2. Spatial particle swarm optimization model for optimal allocation of space utilization.

**3.2.1. Factors affecting the use of building space.** The distribution of architectural space is affected by various factors, and the degree of interaction varies between different regions and varieties. By sorting out the data related to land and buildings in various regions, the weight coefficients of various indicators affecting the use of building space are summarized, and the main impact indicators are selected according to the size of the values. As each indicator has different quantity units and magnitudes, it needs to be standardized. The author adopts the maximum standardization method to normalize the data, as shown in Formula (3.1).

$$D_{ij} = 100 \times d_{ij} / \max(d_i) \quad (3.1)$$

In the formula,  $D_{ij}$  represents the normalized value of the  $j$ th data of the  $i$ th index affecting the utilization and allocation of building space;  $f_{ij}$  represents the true value of the  $j$ th data of the  $i$ th index affecting the utilization and allocation of building space;  $\max(d_i)$  represents the maximum value in the  $i$ th index [2, 16].

The following main indicators can be obtained after data normalization by formula (3.1): Total amount of land development, urban construction, rural land, industrial and mining land, commercial land, and green land. The above indicators will affect the utilization and configuration of building space to varying degrees, and also affect each other, the specific expression is Formula (3.2).

$$Q_i = af(x_i) + b \quad (3.2)$$

In formula (3.2),  $Q_i$  represents the land use index mentioned above;  $x_i$  refers to the factors that affect the utilization and allocation of building space, and  $a$  and  $b$  are weight coefficients and adjustment factors [5].

**3.2.2. Objective function of space optimal configuration.** In order to improve the utilization rate of building space, it is necessary to optimize the configuration of building space. The objective function is used to evaluate the benefits of the building space configuration scheme, and the accuracy of its expression directly affects the effectiveness of the optimized structure. The author constructs the objective function of optimal allocation of architectural space from economic, social and ecological benefits. The economic benefits are used to evaluate the economic benefits generated by the configuration scheme of the building cluster, which can be quantitatively described by monetized indicators. The author uses the ratio of GDP and building area produced by each building area to characterize the economic benefits of the buildings in this area. However, the total GDP is also related to the degree of economic investment in the region. Therefore, it is necessary to comprehensively consider economic benefits from multiple factors such as land use area, total investment and GDP. Social benefits are used to characterize the ability of buildings in this area to meet people's daily life and work needs, and pay more attention to whether the configuration of building clusters is conducive to social equity and the improvement of people's quality of life. The indicators involved mainly include: Land utilization rate, urbanization rate, traffic network structure index, coverage rate of social service facilities, etc. Compared with economic and social benefits, ecological efficiency pays more attention to the impact of the configuration of building clusters on the ecological environment. In order to meet the economic development mode of "resource saving and environment-friendly", the impact of land use and building configuration on land degradation and soil pollution should be minimized. The author uses forest land coverage, wetland coverage and per capita green land coverage to evaluate the ecological benefits [6, 9].

Assume that the building space utilization rate of each area in a city is represented by  $E_1, E_2, E_3, \dots, E_n$  according to its value. The total building space is expressed in  $L$ , and the total building space allocated to each area is  $L_1, L_2, L_3, \dots, L_n$ . Therefore, the linear programming model used to plan the optimal allocation of building space utilization in the city can be expressed by equations (3.3) and (3.4).

$$E_i = E(x_i, L_i) \quad (3.3)$$

$$L = L_1, L_2, \dots, L_n \quad (3.4)$$

In the formula,  $x_i$  represents the exogenous element of non construction land. Through the analysis of the author, it can be seen that the realization of the optimal allocation of the utilization of building space is achieved by the cooperation of multiple objectives [12, 14].

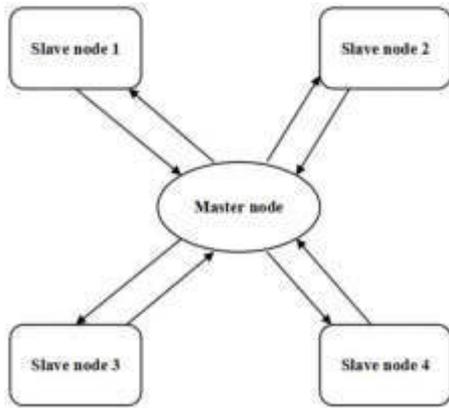


Fig. 3.3: Schematic diagram of master-slave parallel algorithm structure

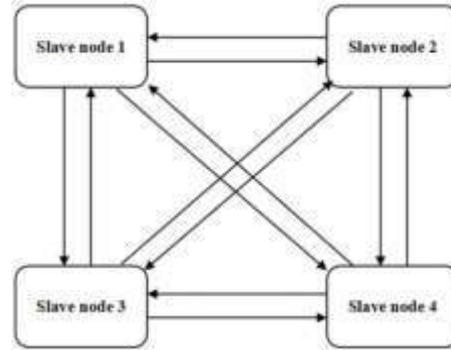


Fig. 3.4: Structure of point-to-point parallel algorithm

**3.2.3. Parallel space utilization optimization configuration model.** Most of the building space utilization and configuration algorithms are based on serial algorithms, which can not satisfy the needs of large area and high precision. Meanwhile, the serial data transfer takes a long time, which makes it hard to guarantee the timeliness. A parallel algorithm is used to improve the performance of PSO. Parallel algorithms are usually divided into master-slave and point-to-point. The author analyses their merits and demerits and makes improvements on them [20, 10]. In order to meet the requirements of large area, high precision and low time consumption, the author studied the parallelization algorithm of optimal allocation of building space utilization. All the particles are divided into different subsets by sub-clustering, and a parallel particle swarm optimization model is developed according to the corresponding communication strategy.

For the optimal configuration of building space utilization, the general idea of the particle swarm optimization algorithm is to divide the overall optimization configuration task into multiple sub tasks according to the region for simultaneous implementation, that is, the particle swarm is divided into multiple sub populations, and each sub group can perform the optimization task independently. In the process of independent optimization, each sub group will communicate in time to share the best configuration scheme location. In order to reduce the cost and time consumption, the author uses the coarse grain parallel model to build the optimal allocation model of building space utilization [18]. The master-slave parallel algorithm framework is shown in Figure 3.3. The algorithm includes master node and slave node. When there is a task, the master node will decompose and distribute the task to each slave node, and communicate with the slave node in a timely manner. The slave node is responsible for completing the assigned tasks and maintaining communication with the master node. Because the master node does not participate in the task execution, the parallel efficiency is wasted. At the same time, the slave nodes cannot communicate directly, which increases the communication workload of the master node, and the communication efficiency is low. The point-to-point parallel algorithm framework is shown in Figure 3.4. This form discards the distinction between master and slave nodes, and each node can directly communicate with each other, which improves the communication efficiency. However, when the number of nodes is large, the communication network is complex. According to the above two parallel forms, the author proposes a chain parallel structure, as shown in Figure 3.5. The chained parallel architecture has multiple master nodes to allocate tasks, while maintaining communication with slave nodes and other master nodes. Slave nodes belonging to the same master node can directly communicate with each other to improve communication efficiency [7, 17].

#### 4. Result Analysis.

**4.1. Simulation test and verification.** In view of whether the building space utilization optimization algorithm based on spatialized particle swarm optimization proposed by the author can be applied to the building space configuration problem of multiple regional scales, the author uses three regions to test and verify.

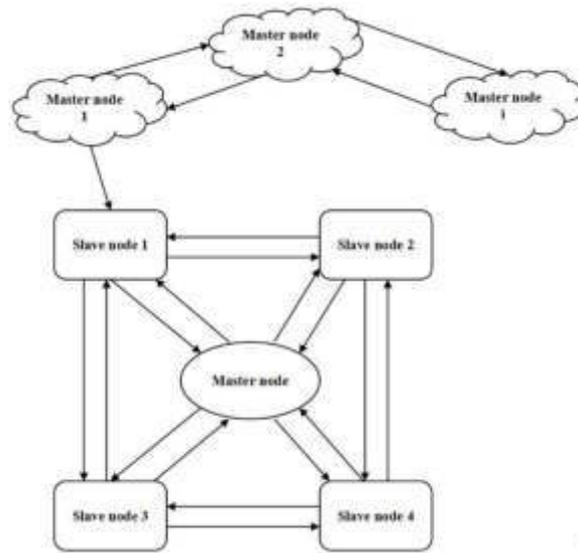


Fig. 3.5: Structure of chained parallel algorithm

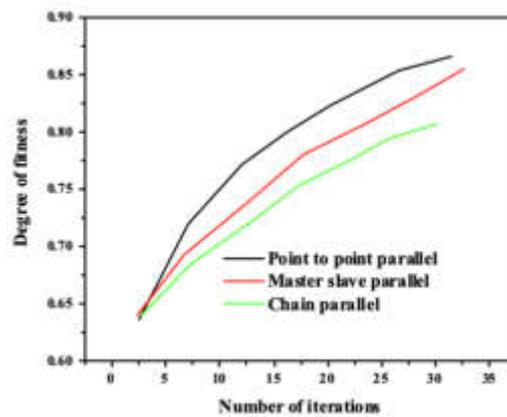


Fig. 4.1: Fitness comparison of three parallel models with different iterations

The experimental data is based on the utilization of building space in these three regions in 2015, the vector map is divided into  $30\text{ m} \times 30\text{ m}$  sized grids, and all statistical data and spatial data are projected on each grid cell. The author tested the convergence speed and running time of the spatialized particle swarm optimization parallel algorithm.

To verify the validity of the proposed chain parallel architecture, we compare it with the master-slave parallel structure and the point to point parallel structure [15, 19]. Figure 4.1 shows the fitness convergence curves of the three parallel models under five process parallelism. It can be seen from Figure 4.1 that the difference between the fitness values of the final convergence of the three parallel models is small, and the main difference is the convergence speed. The convergence speed of the chained parallel model proposed by the author is obviously better than that of the master-slave model and the point-to-point model.

During the run time test, set the three parallel models to run under the conditions of 8, 16, 32 and 64 nodes respectively, and the results are shown in Table 4.1 [8, 1]. Because of the combination of the advantages of master-slave model and point-to-point model, the running time of chained parallel model is significantly lower

Table 4.1: Running time of three parallel models/s

Model type	8 nodes	16 nodes	32 nodes	64 nodes
Chain type	143	86	43	31
Master-slave	184	112	76	47
Point to point	156	98	78	93

than that of the other two parallel models. Especially when the number of nodes is large, the setting of multiple master nodes can allocate tasks with greater degrees of freedom without increasing too much communication burden.

**5. Conclusion.** The optimum model of indoor space utilization was established by PSO. Based on the number of buildings and the total area, this paper makes a general improvement on the PSO algorithm, and defines some relative optimization strategies. The symbol coding method is used for spatial coding, the maximum standardization method is used for data processing, and the influencing factors of space utilization are summarized to obtain the objective function of the optimal configuration of building space. The advantages and disadvantages of master-slave parallel model and point-to-point parallel model are analyzed, and the optimization algorithm of chain parallel structure is proposed. The data simulation test results show that the optimization algorithm proposed by the author has good performance and can meet the conventional optimization requirements of building space utilization.

## REFERENCES

- [1] A. H. AL KA'BI, *Comparison of energy simulation applications used in green building*, Annals of Telecommunications, 75 (2020), pp. 271–290.
- [2] D. CARRIÓN, E. COLICINO, N. F. PEDRETTI, K. B. ARFER, J. RUSH, N. DEFELICE, AND A. C. JUST, *Neighborhood-level disparities and subway utilization during the covid-19 pandemic in new york city*, Nature communications, 12 (2021), p. 3692.
- [3] D. R. X. DE OLIVEIRA, G. MOREIRA, A. R. DUARTE, A. CAÑADO, AND E. LUZ, *Spatial cluster analysis using particle swarm optimization and dispersion function*, Communications in Statistics-Simulation and Computation, 50 (2021), pp. 2368–2385.
- [4] M. GHAHRAMANI, A. O'HAGAN, M. ZHOU, AND J. SWEENEY, *Intelligent geodemographic clustering based on neural network and particle swarm optimization*, IEEE Transactions on Systems, Man, and Cybernetics: Systems, 52 (2021), pp. 3746–3756.
- [5] M. JANGRA AND B. SINGH, *Mod-k-chained variant of present and clefia lightweight block cipher for an improved security in internet of things*, SN Computer Science, 3 (2022), pp. 1–15.
- [6] S. KALIKAR, C. JAIN, M. VASIMUDDIN, AND S. MISRA, *Accelerating minimap2 for long-read sequencing applications on modern cpus*, Nature Computational Science, 2 (2022), pp. 78–83.
- [7] J. LATT, O. MALASPINAS, D. KONTAXAKIS, A. PARMIGIANI, D. LAGRAVA, F. BROGI, M. B. BELGACEM, Y. THORIMBERT, S. LECLAIRE, S. LI, ET AL., *Palabos: parallel lattice boltzmann solver*, Computers & Mathematics with Applications, 81 (2021), pp. 334–350.
- [8] T. LI, L.-M. YUAN, G.-Q. HOU, AND Y.-F. WU, *Rapid design and construction management of emergency hospital during the covid-19 epidemic*, Structural Engineering International, 32 (2022), pp. 142–146.
- [9] I.-C. LIN, Y.-H. YEH, AND K. C.-J. LIN, *Toward optimal partial parallelization for service function chaining*, IEEE/ACM Transactions on Networking, 29 (2021), pp. 2033–2044.
- [10] P. PIROZMAND, H. ALREZAAMIRI, A. EBRAHIMNEJAD, AND H. MOTAMENI, *A new model of parallel particle swarm optimization algorithm for solving numerical problems*, Malaysian Journal of Computer Science, 34 (2021), pp. 389–407.
- [11] M. QIN, *Evolution of labor relations in the development of human resources based on improved genetic algorithm*, Journal of Circuits, Systems and Computers, 31 (2022), p. 2250272.
- [12] D. SAMANTA, A. H. ALAHMADI, M. KARTHIKEYAN, M. Z. KHAN, A. BANERJEE, G. K. DALAPATI, AND S. RAMAKRISHNA, *Cipher block chaining support vector machine for secured decentralized cloud enabled intelligent iot architecture*, IEEE Access, 9 (2021), pp. 98013–98025.
- [13] M. SONG, S. LIU, W. LI, S. CHEN, W. LI, K. ZHANG, D. YU, L. LIU, AND X. WANG, *A continuous space location model and a particle swarm optimization-based heuristic algorithm for maximizing the allocation of ocean-moored buoys*, IEEE Access, 9 (2021), pp. 32249–32262.
- [14] C.-W. TSAI, Y.-P. CHEN, T.-C. TANG, AND Y.-C. LUO, *An efficient parallel machine learning-based blockchain framework*, Ict Express, 7 (2021), pp. 300–307.

- [15] H. J. WAGNER, M. ALVAREZ, A. GROENEWOLT, AND A. MENGES, *Towards digital automation flexibility in large-scale timber construction: integrative robotic prefabrication and co-design of the buga wood pavilion*, *Construction Robotics*, 4 (2020), pp. 187–204.
- [16] L. WANG, S. FANG, X. MENG, AND R. LI, *Building extraction with vision transformer*, *IEEE Transactions on Geoscience and Remote Sensing*, 60 (2022), pp. 1–11.
- [17] Y. XIE, M. HE, T. MA, AND W. TIAN, *Optimal distributed parallel algorithms for deep learning framework tensorflow*, *Applied Intelligence*, (2022), pp. 1–21.
- [18] F. YIN AND F. SHI, *A comparative survey of big data computing and hpc: From a parallel programming model to a cluster architecture*, *International Journal of Parallel Programming*, 50 (2022), pp. 27–64.
- [19] R. YUAN, F. GUO, Y. QIAN, B. CHENG, J. LI, X. TANG, AND X. PENG, *A system dynamic model for simulating the potential of prefabrication on construction waste reduction*, *Environmental Science and Pollution Research*, 29 (2022), pp. 12589–12600.
- [20] A. ZAINAB, A. GHAYEB, H. ABU-RUB, S. S. REFAAT, AND O. BOUHALI, *Distributed tree-based machine learning for short-term load forecasting with apache spark*, *IEEE Access*, 9 (2021), pp. 57372–57384.

*Edited by:* B. Nagaraj M.E.

*Special issue on:* Deep Learning-Based Advanced Research Trends in Scalable Computing

*Received:* Aug 3, 2023

*Accepted:* Oct 17, 2023



## OPTIMIZATION OF NONLINEAR CONVOLUTIONAL NEURAL NETWORKS BASED ON IMPROVED CHAMELEON GROUP ALGORITHM

QINGTAO ZHANG\*

**Abstract.** In order to solve the most difficult problem of the architectural model established by CNN in solving specific problems, which results in parameter overflow and inefficient training, an optimization algorithm for nonlinear convolutional neural networks based on improved chameleon swarm algorithm is proposed. This article mainly introduces the use of Chameleon Swarm Optimization (PSO) algorithm to research the parameters of CNN architecture, solve them, and achieve the optimization of the optimization model. Although the number of parameters that need to be set up in CNN is very large, this method can find better testing space for Alexnet samples with 5 different images. In order to improve the performance of the improved pruning algorithms, two candidate pruning algorithms are also proposed. The experimental results show that compared with the traditional Alexnet model, the improved pruning method improves the image recognition ability of the Caffe primary parameter set from 1.3% to 5.7%. Conclusion: This method has wide applicability and can be applied to most neural networks which do not require any special functional modules of the Alexnet network model.

**Key words:** Deep learning, Convolutional neural network, Chameleon group optimization algorithm, Image recognition

**1. Introduction.** The optimization problem, which dates back to the ancient extremum problem, is a branch of computational science and is now a widely studied topic. Optimization problem requires that the maximum or minimum value of the objective function can be obtained by reasonable search method under certain constraints. Since the target space of the optimization problem is generally huge, it is impossible to use the exhaustive method to solve it. It is necessary to design a suitable optimization method to solve it.

Traditional optimization methods include simple form method, common gradient method, Newton method, etc. Because these optimization methods generally require the objective function to be differentiable, and need to search the search space on a large scale, the algorithm is feasible in theory but not in practice. However, the actual optimization problems are often complex, with the characteristics of non-differentiable, nonlinear and multi-extremum. It is obvious that the traditional optimization method can not meet the requirements of calculation accuracy and convergence speed. Therefore, designing efficient algorithms to solve complex optimization problems has always been a research hotspot in computational science. Optimization problem has always been an important problem to be solved in the field of scientific research and engineering. It has played an important role in the development of the history of science and the progress of human civilization, which makes the study of optimization theory become a very active field. With the deepening of human's understanding and research of the natural world, the scale of the problems involved is also growing, and the large-scale optimization problem becomes an urgent problem to be solved effectively, which raises the requirement for optimization theory. Some optimization theories and algorithms have been further developed due to the increasing performance requirements and the continuous improvement of computing performance. At the same time, these optimization technologies have been successfully applied to a number of practical engineering fields, and achieved certain development, such as industrial production control task scheduling intelligent system [13, 5].

Convolutional neural network is a deep pre crushing neural network, whose basic characteristics are obtained by continuous clustering of extraction methods. It is composed of an access layer, a hidden layer, an entire link layer, and an output layer, which are connected by a solution layer and a sub-layer.

The convolution layer deals with the resolution function of the object image, and extracts the feature of the image. The weight is the same as the window, first in the horizontal translation, then the bottom, so the

---

\*Department of Computer and Information Engineering, Hebei Petroleum University of Technology, 067000, China (QingtaoZhang7@163.com)

operation of the picture is the operation resolution.

In recent years, deep neural networks have become a phenomenon research hotspot due to their remarkable performance advantages, and have been applied to various fields and made remarkable achievements. Deep neural network has greatly exceeded the traditional algorithm architecture in terms of performance, namely, the way of manual features and classifiers, and has been favored by experts and scholars in various fields. The development of neural network began in the late 19th century, and it has been more than a century. From the original MP model to the perceptron model; However, in the trough period, because the feasibility of using multi-layer structure to extend the perceptron model cannot be verified theoretically, the neural network shows certain limitations when dealing with nonlinear problems.

In a sense, the training process of neural network is also the process of dealing with large-scale optimization problems, that is, looking for network parameters that make the model adapt to the data. The mainstream approach to deal with optimization problems on neural network is to use error back propagation (Error Back Propagation, BP). In addition, parameters are updated in the form of error gradient descent. This traditional optimization method needs to calculate the gradient of each parameter in the face of the huge number of parameters in the deep network, which increases the difficulty of solving and requires very high computing power, and the current equipment cannot meet its fast solution.

**2. Literature Review.** Convolutional neural network (CNNs), as one of the most important depth models, has good feature extraction and generalization ability. It has achieved great success in image processing, target tracking and detection, natural language processing, scene classification, face recognition, audio retrieval, medical diagnosis and many other fields. On the one hand, the rapid development of high-resolution neural network is due to the significant improvement in computer performance, which makes the construction and training of large-scale networks not restricted by the hardware level. On the other hand, due to the development of large-scale data processing, the general scalability of the network is enhanced.

Convolutional neural networks (CNN) is an important method in image recognition, including image resolution technique, clustering technique, and composite layers. Pop, C. B et al. puts forward a model of AlexNet, which is the first time that CNN is superior to the traditional mathematics model. Based on the LeNet-5 model, it is proposed to extend and deepen the network, and to improve the recognition capability of the model. These ideas have received the approval of the scientists and the CNNs with complex, multi - and multi - constraints [14].

Appropriately increasing the scale of network model and training data is helpful to improve the final recognition effect of neural network, but it is bound to be accompanied by a huge amount of computation and long training time. Therefore, the acceleration of convolutional neural networks is now the focus of research. The main acceleration modes focus on the adjustment training algorithm and parallel acceleration. Parallel acceleration mainly uses hardware environment and parallel computing framework to accelerate hardware. FPGA implementation of convolutional neural network has appeared as early as the mid-1990s, which uses arithmetic methods with low accuracy to replace all multipliers. In recent years, more and more studies have been conducted on using FPGA to accelerate convolutional neural network. In addition, Bell LABS implemented ANNA's chip in the early 1990s, which was also the first time to accelerate convolutional neural networks through hardware. In recent years, the Institute of Computing Science of the Chinese Academy of Sciences has proposed deep learning processors DianNao Computer and DaDianNao large computer. Aiming at the underlying hardware of the computer, the operations of each layer of deep learning are integrated into a hardware unit, which can well improve the efficiency of convolutional neural networks [9].

The CNN model proposed in this paper through parameter optimization of the improved chameleon group algorithm has better recognition accuracy than the standard CNN model and can be verified. The method proposed in this paper is suitable for most existing CNN architectures [12].

### 3. Research Method.

#### 3.1. Chameleon algorithm.

**3.1.1. The main idea of the algorithm Chameleon.** Chameleon is a hierarchical clustering algorithm that uses qualitative models. In the Chameleon clustering method, two clusters are merged if the intersections and computations between them correspond to the intersections and computations of items in the cluster. The

way it works is to put the data items into many small sub-clusters from a shared image, and then combine the sub-classes with a hierarchical clustering algorithm to see the actual results. A unified process model helps detect natural or homogeneous groups and can be applied to any type of data as long as the features are similar. The Chameleon algorithm takes into account cluster connectivity and computation, especially the intrinsic properties of clusters, to identify similar subclusters [10, 18].

**3.1.2. Chameleon algorithm.** The Chameleon algorithm defines its properties as a k-nearest neighbor graph. Each K-point in the nearest neighbor graph represents a data object, and if data A is one of the k-closest objects of data B, then objects A and B are edges. The nearest image K-concept is dynamically obtained. Community: The electrical community of an object is determined by the density of its siblings. The idea of K-community is expressed dynamically: the local electricity of an object is determined by the density of the place where the object is. In densely populated areas, the definition of community is narrow. In the distribution of objects section, the definition of groups is broader and the area density is denoted as edge weight. Therefore, the edges of a dense object have more weight than the edges of a diffuse object.

**3.1.3. The determination of similarity between clusters in the Chameleon algorithm.** The Chameleon determines the similarity between clusters by the relative interconnection  $RI(C_i, C_j)$  and the relative approximation  $RC(C_i, C_j)$  of two clusters. Chameleon.

(1) Relative interconnection  $RI(C_i, C_j)$  is defined as the absolute interconnection between  $C_i$  and  $C_j$  and the normalization of the internal interconnection of two clusters, i.e., the following formula (3.1):

$$RI(C_i, C_j) = \frac{|E_{C_i, C_j}|}{\frac{1}{2}(|E_{C_i}| + |E_{C_j}|)} \quad (3.1)$$

$E_{C_i, C_j}$  is the truncated edge of the cluster containing  $C_i$  and  $C_j$  classified into  $C_i$  and  $C_j$ ;  $E_{C_i}$  (or  $E_{C_j}$ ) is the size of the minimum truncated bisector (that is, the weighted sum of the edges that need to be cut off to divide the graph into two roughly equal parts)

(2) Relative approximation  $RC(C_i, C_j)$  is defined as the normalization of the absolute approximation between  $C_i$  and  $C_j$  about the internal approximation of the two clusters, namely, the following equation (3.2):

$$RC(C_i, C_j) = \frac{S_{EC(C_i, C_j)}}{\frac{|C_i|}{|C_i| + |C_j|} S_{EC} + \frac{|C_j|}{|C_i| + |C_j|} S_{EC_{C_j}}} \quad (3.2)$$

$S_{EC(C_i, C_j)}$  is the average weight of the edges connecting vertices  $C_i$  and  $C_j$ ,  $S_{EC}$ ,  $S_{EC_{C_j}}$  are the average weights of the edges of the minimum truncated bisector of  $C_i$  and  $C_j$ , respectively.

**3.2. Parameter to be optimized.** Related issues that need to be optimized when solving neural network problems are the size of the convolutional kernel and the size and type of weighting layers for each feature parameter. convolutional windows network. Parameters are looked up using high-valued floating-point values, then balanced, keeping the desired number of objects and taking into account the parameters to fall back to the range if they are lost outside of the dynamic configuration. This is because if the step size is optimized by the improved chameleon swarm algorithm, the size of the image to be processed will be very small, and the method of extracting local features from the solution will not work well. In this study, no steps have been taken during the resolution to allow a larger area to search for other constraints, and a buffer pool [6, 1].

The introduction of nonlinear function theory is mainly to improve the teaching ability of the network and to make deep neural multi- points.

If it is necessary to improve the parameters related to the nonlinear network, for example, the number of network layers, the GA can get a better effect, the effect of the modified PLOS is that the specified length is fixed, so it is not suitable for the dynamic swarm model [17].

**3.3. Optimized process design.** The flow chart of CNN optimization by using the improved chameleon swarm algorithm proposed in this paper is shown in Figure 3.1, where Y indicates that conditions are met, while N indicates that conditions are not met.

In this study, the learning algorithm of the algorithm neural network is calculated as a particle swarm algorithm to construct the chameleon swarm algorithm. Therefore, the number of process particle is the

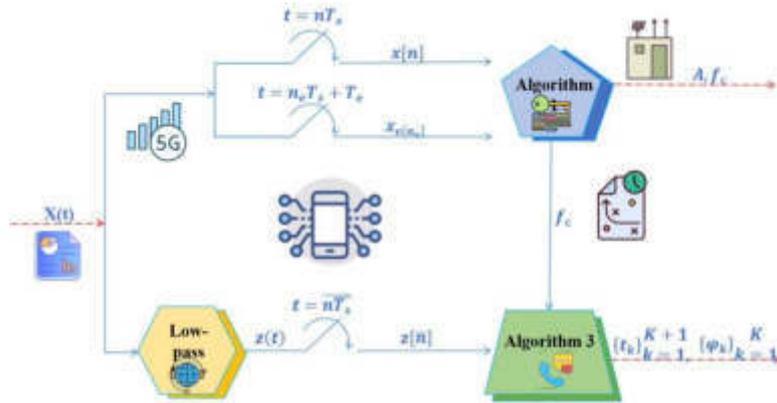


Fig. 3.1: Improved Chameleon Swarm Algorithm -CNN Training flowchart

number of network courses. First, start the task and speed of this product, then calculate the fitness function according to the error of actual result and requirement, and use the world look good and speed of each bit to get the weight of the network. Those. New weights are then substituted and iterated so that the algorithm stops until the fitness changes to some threshold [4, 15].

**3.4. Optimization and improvement.** Ideally, the performance of each sub-scale should be assessed in the same way as the last stage of training, the same duration, the same number of training sessions and so on. But this is not true, because if the number of particles is  $M$  and the number of epochs optimized by the chameleon swarm is  $N$ , then the optimum time of this parameter is  $MN$ . But there are too many optimization parameters in ANN, which results in a long training time and a high cost. In addition, many regular training in the database can lead to excessive interference, resulting in a number of performance issues in practical applications. It should be noted that during the optimization process, there is no need to know a small proportion of the values of each parameter for optimal performance. Only the value of the parameter is superior to other parameters, which means that its function is the best solution for particle operation. Therefore, this article proposes two development methods to improve efficiency. When CNN control realizes capability, it reduces the time parameter optimization and shortens the training time.

Dependency classification model: Parameter resolution is dependent on data size and quality. That is to say, when interpreting the data, we can estimate the number of iterations required to build an optimum neural network. As mentioned above, this article first introduces CNN network frequency, then selects the object as focus, and finally calculates Spearman correlation coefficient of particles according to the calculation results. The formula is as follows: (3.3) (3.4):

$$\rho = 1 - \frac{6 \sum_{i=1}^n (d_i)^2}{n_0} \quad (3.3)$$

$$d_i = \sigma(p_i) - \sigma(q_i), n_0 = n^3 - n \quad (3.4)$$

Based on the correlation coefficient, the predetermined epoch number threshold value  $E$  can be obtained, which can be used as the training turn value of CNN optimization by using the improved chameleon swarm algorithm [20, 8].

Process-based transformation: Although the results obtained from the correlation level based method are reliable, this process requires extensive training of CNN, which takes a lot of manpower resources. Therefore, this article presents another change based approach. Because of the randomness of backpropagation during initial training phase, the performance accuracy of the network structure is unstable. However, as training increases, the number of training cycles on data gradually increases, and the recognition accuracy often results

Table 4.1: The amount of data in the data set used for the experiment

	Train	Test	Val
CIFAR10	50000	10000	N/A
CIFAR100	50000	10000	N/A
Subset10	12081	1500	500
Subset20	37476	4500	1500
Subset50	59907	7500	2500

Table 4.2: Parameters to be optimized and their ranges

layer	hyper-parameter	dynamic range
Convolutional Layer	Number of feature maps	50~180
	Padding size	0~7
	Convolution kernel size	2~7
Pooling Layer	Pooling type	MAX, AVE
	Pool core size	2~7

in more stable performance, appearing in network structure Therefore, the stability of the network structure can be expressed as follows, shown in formula (3.5):

$$CV = \frac{\mu}{\sigma} \quad (3.5)$$

Here is the formula (3.6) (3.7):

$$\mu = \frac{\sum_i^N \text{accuracy}^k[i]}{N} \quad (3.6)$$

$$\sigma = \sqrt{\frac{\sum_i^N (\text{accuracy}^k[k] - \mu)^2}{N}} \quad (3.7)$$

$\text{accuracy}^k[k]$  is the classification accuracy of the  $i$  th particle in the  $k$  stage. Allowing comparison of different performance of the network structure in different cycles. Therefore, when the electric field is unstable, the best accuracy for the particle size can be obtained by comparison. This method will not require an optimal algorithm for the final classification, thus reducing the computation cost [3].

#### 4. Interpretation of Result.

**4.1. Experimental design.** The data used in this article includes CIFAR10, CIFAR100, and ImageNet data, which can be divided into 10 classes, 20 classes, and 50 classes. respectively. Table 4.1 shows how many images are used for training and testing across all documents. In this paper, we apply Alexnet Network Model and Improved Chameleon Swarm Algorithm to optimize the data classification efficiency. In this paper, we use the simplified model of Alexnet to improve the precision of classification.

Parameters that need to be optimized in the training stage are shown in Table 4.2, and there are about  $3.6 \times 1020$  possibilities for parameter setting. Therefore, even for the standard Alexnet model configuration, a simple and direct search is not possible. In this experiment, the hyperparameter of the improved chameleon swarm algorithm is set as  $c_{r1} = c_{r2} = 1.494, \omega = 0.792$ .

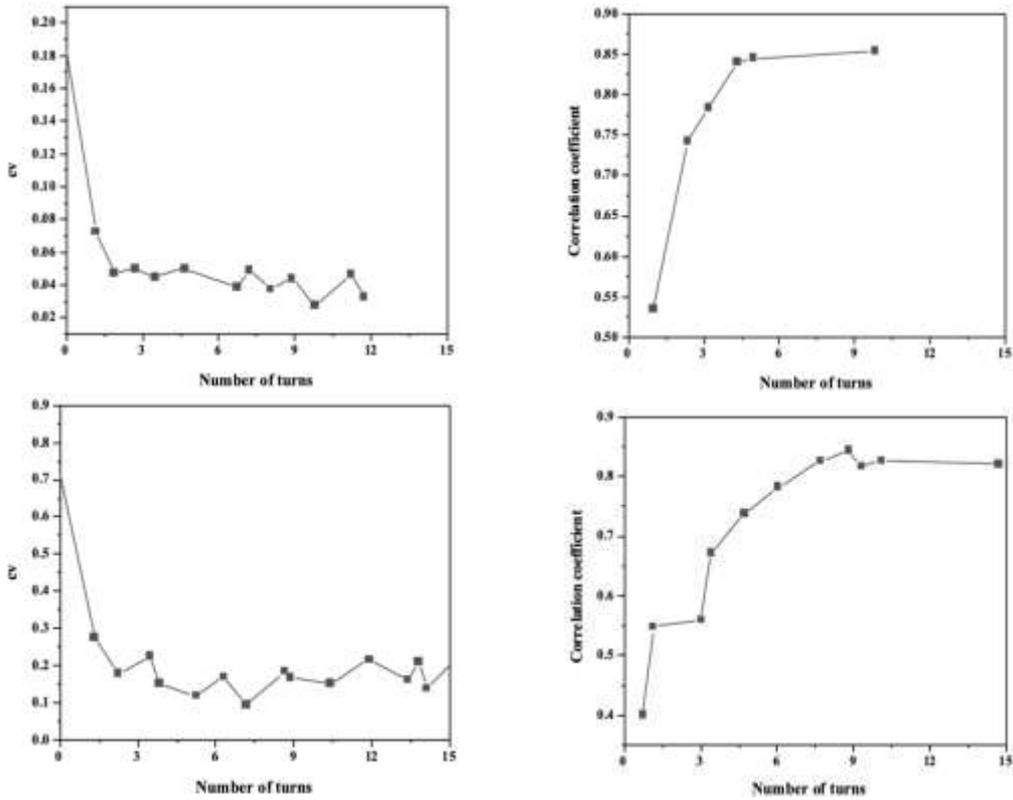


Fig. 4.1: Chart of variation of correlation coefficient between volatility and ranking

**4.2. Analysis of experimental results.** In this paper, the CNN model shown in Figure 4.1 (a) (b) (c) (d) is related to erêmbîn and the chameleon group algorithm is developed and optimized based on transformation. As can be seen from the figure, the relationship level will be above 0.8 and the change will gradually stabilize. However, evolutionary based technique does not require more quantitative analysis of data for comparison and better for optimization of neural network. As shown in Figure 4.1, in the next test, the number of training sequences for the CIFAR10 data and CIFAR100 data will be set to 5 and 10, respectively. The number of training sequences for the CIFAR10 data and CIFAR100 data will be set to 5 and 10, respectively. By limiting the number of training rounds, the load ratio of the improved chameleon swarm algorithm can be reduced as compared to the original improved chameleon swarm algorithm [2].

On the basis of improved chameleon clustering algorithm, this article studies the relationship between classification accuracy and iteration number of CNN model by parameter optimization. The number of optimization methods in this model is 21. Considering the training cost, the first number of bits per unit is 15, and the number of iterations is 0-60. Best and Average corresponds to the best and average of 15 sections. This figure confirms that the performance of the Alexnet network model improves as the iteration time increases. However, it should be noted that the development of the chameleon swarm algorithm for random initialization only approximates the network model to the optimal model and cannot guarantee that the best view of the world has been found [16, 7].

In Table 4.3, the image classification performance of Alexnet model optimization using improved chameleon swarm algorithm was compared with that of Alexnet model. In different data tested, the classification accuracy of network model was improved using improved chameleon swarm algorithm was superior to different level, but it was only 2% to 4% higher than the BP method usually used in the past. At the same time, it also shows that, because the improved chameleon swarm algorithm is essentially an optimization process based on random

Table 4.3: Improved Chameleon Swarm Algorithm -Alexnet performance compared to Standard Alexnet

Data set	Alexnet	Improved chameleon group algorithm-Alexnet	Performance difference	optimal cost
CIFAR10	77.76%	80.25%	2.48%	2%
CIFAR100	52.43%	55.67%	3.24%	3%
Subset10	72.57%	74.83%	2.26%	3%
Subset20	59.69%	65.46%	5.78%	4%
Subset50	57.56%	58.85%	1.29%	4%

allocation, no data set is necessarily the best, but the training results that are relatively close to the essential effect of the model can be found in a large number of training. Moreover, the performance of the model can be improved continuously by using the modified PGA [11, 19]. The convergence of the Chameleon algorithm is demonstrated. It is proved that the Chameleon Algorithm can converge to the global optimum with the increase of time.

**5. Conclusion.** This paper proposes a particle-particle optimization algorithm to optimize neural network constraints to solve neural network problems. Local minimum value due to multiple parameters. In this paper, the parameter setting and function selection of convolutional neural networks are explored and experimented, and the influence of these parameters on model training is revealed. The experimental results show that by pre-setting the relevant hyperparameters and matching with RMSProp or Adam, the accuracy of convergence can be reached faster, thus improving the training efficiency of convolutional neural networks. In the course of design, this article discusses how to reduce the number of computations by optimizing the data and controlling the number of training courses, in order to make the experience more satisfactory. This article tries to prove that by improving the chameleon swarm algorithm, the image recognition accuracy of the improved Alexnet model is 1.3 to 5.7 times higher than that of the traditional online training model. Meanwhile, the model proposed in this article is independent of the specific structure of Alexnet network model, so this model is universal and can be applied to most neural networks.

## REFERENCES

- [1] A. ALI, Y. ZHU, AND M. ZAKARYA, *Exploiting dynamic spatio-temporal graph convolutional neural networks for citywide traffic flows prediction*, Neural Networks, 145 (2022), pp. 233–247.
- [2] A. ARABALI, M. KHAJEHZADEH, S. KEAWSAWASVONG, A. H. MOHAMMED, AND B. KHAN, *An adaptive tunicate swarm algorithm for optimization of shallow foundation*, IEEE Access, 10 (2022), pp. 39204–39219.
- [3] M. S. BRAIK, M. A. AWADALLAH, M. A. AL-BETAR, A. I. HAMMOURI, AND R. A. ZITAR, *A non-convex economic load dispatch problem using chameleon swarm algorithm with roulette wheel and levy flight methods*, Applied Intelligence, (2023), pp. 1–40.
- [4] Y. DONG, Q. LIU, B. DU, AND L. ZHANG, *Weighted feature fusion of convolutional neural network and graph attention network for hyperspectral image classification*, IEEE Transactions on Image Processing, 31 (2022), pp. 1559–1572.
- [5] M. GHASEMI, M.-A. AKBARI, C. JUN, S. M. BATENI, M. ZARE, A. ZAHEDI, H.-T. PAI, S. S. BAND, M. MOSLEHPOUR, AND K.-W. CHAU, *Circulatory system based optimization (csbo): An expert multilevel biologically inspired meta-heuristic algorithm*, Engineering Applications of Computational Fluid Mechanics, 16 (2022), pp. 1483–1525.
- [6] T. GHAZAL, *Convolutional neural network based intelligent handwritten document recognition*, Computers, Materials & Continua, 70 (2022), pp. 4563–4581.
- [7] G. HABIB AND S. QURESHI, *Optimization and acceleration of convolutional neural networks: A survey*, Journal of King Saud University-Computer and Information Sciences, 34 (2022), pp. 4244–4268.
- [8] G. HU, R. YANG, AND G. WEI, *Hybrid chameleon swarm algorithm with multi-strategy: A case study of degree reduction for disk wang-ball curves*, Mathematics and Computers in Simulation, 206 (2023), pp. 709–769.
- [9] S. IRENE D, J. R. BEULAH, A. K, AND K. K, *An efficient covid-19 detection from ct images using ensemble support vector machine with ludo game-based swarm optimisation*, Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization, 10 (2022), pp. 675–686.
- [10] N. JIA, Y. CHENG, Y. LIU, AND Y. TIAN, *Intelligent fault diagnosis of rotating machines based on wavelet time-frequency diagram and optimized stacked denoising auto-encoder*, IEEE Sensors Journal, 22 (2022), pp. 17139–17150.

- [11] X. KAN, Y. FAN, Z. FANG, L. CAO, N. N. XIONG, D. YANG, AND X. LI, *A novel iot network intrusion detection approach based on adaptive particle swarm optimization convolutional neural network*, Information Sciences, 568 (2021), pp. 147–162.
- [12] G. LIU AND W. MA, *A quantum artificial neural network for stock closing price prediction*, Information Sciences, 598 (2022), pp. 75–85.
- [13] M. NOROOZI, H. MOHAMMADI, E. EFATINASAB, A. LASHGARI, M. ESLAMI, AND B. KHAN, *Golden search optimization algorithm*, IEEE Access, 10 (2022), pp. 37515–37532.
- [14] C. B. POP, T. CIOARA, I. ANGHEL, M. ANTAL, V. R. CHIFU, C. ANTAL, AND I. SALOMIE, *Review of bio-inspired optimization applications in renewable-powered smart grids: Emerging population-based metaheuristics*, Energy Reports, 8 (2022), pp. 11769–11798.
- [15] R. M. RIZK-ALLAH, M. A. EL-HAMEED, AND A. A. EL-FERGANY, *Model parameters extraction of solid oxide fuel cells based on semi-empirical and memory-based chameleon swarm algorithm*, International Journal of Energy Research, 45 (2021), pp. 21435–21450.
- [16] A. SRIDHARAN, *Chameleon swarm optimisation with machine learning based sentiment analysis on sarcasm detection and classification model*, Int Res J Eng Technol, 8 (2021), pp. 821–828.
- [17] C. TIAN, Y. YUAN, S. ZHANG, C.-W. LIN, W. ZUO, AND D. ZHANG, *Image super-resolution with an enhanced group convolutional neural network*, Neural Networks, 153 (2022), pp. 373–385.
- [18] A.-A. TULBURE, A.-A. TULBURE, AND E.-H. DULF, *A review on modern defect detection models using dcnn—deep convolutional neural networks*, Journal of Advanced Research, 35 (2022), pp. 33–48.
- [19] Y. WANG, X. QIAO, AND G.-G. WANG, *Architecture evolution of convolutional neural network using monarch butterfly optimization*, Journal of Ambient Intelligence and Humanized Computing, 14 (2023), pp. 12257–12271.
- [20] J. ZHOU AND Z. XU, *Optimal sizing design and integrated cost-benefit assessment of stand-alone microgrid system with different energy storage employing chameleon swarm algorithm: A rural case in northeast china*, Renewable Energy, 202 (2023), pp. 1110–1137.

*Edited by:* B. Nagaraj M.E.

*Special issue on:* Deep Learning-Based Advanced Research Trends in Scalable Computing

*Received:* Aug 3, 2023

*Accepted:* Nov 3, 2023



## CHANNEL ESTIMATION OF URBAN 5G COMMUNICATION SYSTEM BASED ON IMPROVED PARTICLE SWARM OPTIMIZATION ALGORITHM

XIGANG XIA\*, BO YANG<sup>†</sup> AND ZHIYU LIU<sup>‡</sup>

**Abstract.** In order to solve the problem that the channel estimation accuracy of the traditional urban communication system is not high, the author proposes the channel estimation of the urban 5G communication system based on the improved particle optimization algorithm. This method converts the channel estimate into a regression fit and adjusts the fit. Focusing on regression fitting problems, big data models are used to display offline data, study channel nonlinearities, and obtain initial channel prediction models. To solve the adaptive problem, the author collects real-time teaching data in a real online learning mode and integrates blended learning to update the model, to avoid the problem of overspending on offline training. Offline tests show that the performance of the channel estimation model is the best for different channels. As the signal-to-noise ratio increases, the MSE value is stable at around 1200. Conclusion: The channel estimation method can produce different characteristics of channel estimation in different situations and improve the signal recovery function of the communication system.

**Key words:** Channel estimation, Improved particle swarm optimization algorithm, Integrated learning

**1. Introduction.** With the popularization of smart terminals and the emergence of many new mobile applications, there will be high requirements for wireless transmission speed, minimum delay, and network access support for multiple terminals. It can be assumed that a 4G system will be available in the near future. Due to the difficulty of meeting the demand for mobile communication services, countries around the world have started researching 5th generation (5G) mobile communication. Those. By 2022, the capacity of commercial 5G systems should be 1,000 times greater than the current 4G systems.

In the context of the continuous expansion of the scale of mobile communication services, 5G has become a new generation of mobile communication technology. 5G networks will cost less, consume less energy, have larger communication capacity, and be more secure and reliable. Compared with 3G and 4G, 5G's transmission rate can reach the millisecond level, and the device connection density is 10 to 100 times higher than that of 4G. In addition, 5G network can overcome the space-time limitation of information communication, greatly shorten the distance between people and things, and realize high-speed interconnection between people and things, for example, iot applications such as smart grids, smart healthcare, and driverless cars are expected to become a reality. Compared to 5G, 3G and 4G are more focused on mobile broadband applications, while 5G networks are more focused on the speed of wireless connections. The advantages of 5G are not only reflected in individual users, but also in public safety, for example, remote monitoring of emergency call drones and tracking of emergency personnel rely on the high speed, high density connection and high reliability of 5G networks. The continuous increase in demand for new 5G technologies and the emergence of new communication scenarios have promoted the emergence of technologies with development potential in 5G communication systems.

In particular, the proposal of massive MIMO technology has become a hot topic in the industry. Massive MIMO technology is suitable for multi-user scenarios on the side of 5G base stations. If the number of antennas at the base station side goes to infinity, the user channel vectors gradually tend to be orthogonal. Theoretically, this eliminates user interference in the same cell. The increase in the number of antennas also brings a series of problems and challenges, such as pilot pollution. At present, the theories related to massive MIMO technology have been relatively mature, but more comprehensive and in-depth research is needed to give full play to the role of the technology, and channel estimation and equalization technology is one of the key research directions.

---

\*Jilin JLU Communication Design Institute Co., Ltd, Jilin, China

<sup>†</sup>Jilin JLU Communication Design Institute Co., Ltd, Jilin, China

<sup>‡</sup>Jilin JLU Communication Design Institute Co., Ltd, Jilin, China (Corresponding author)

Multiple-input-multiple-output (MIMO) technology can simultaneously reuse multiple data streams on the spectrum by installing multiple antennas at the transmit and receive sites to improve system spectrum utilization. By installing more antennas at the base station, more data streams can be used simultaneously back on the spectrum device, which means that more terminal users can work at the same time. In order to support multiple UTs simultaneously with massive MIMO systems, MIMO systems with multiple antennas installed at the base station are being advocated in academic and commercial circles. And they do a lot of research on MIMO systems.

At present, channel estimation methods for massive MIMO mainly include blind channel estimation, semi-blind channel estimation, channel estimation, and compressed sensor-based estimation. Among them, the test based on channel estimation is less computationally intensive and easy to implement, such as LS channel estimation and MMSE channel estimation. However, this kind of channel estimation algorithm needs to be supplemented with pilot signal, which is easy to produce pilot pollution, which is the development bottleneck of massive MIMO system. Compared with the traditional LS and MMS channel estimation algorithms, the LMMSE algorithm based on singular value decomposition, the estimation performance is better and the pilot overhead and feedback overhead are less [6].

By installing dozens or hundreds of antennas at the BS end, large-scale MIMO can simultaneously support a large number of antenna users, among which the number of antennas is greater than the number of antennas of the base station receiver  $M$ . users  $K$ . Compared with traditional MIMO systems, MIMO systems can improve system connectivity, spectrum efficiency, and energy efficiency. Therefore, massive MIMO technology is proposed as a candidate for 5G in the future.

**2. Literature Review.** Signal estimation is usually realized by pilot, semi-blind and blind estimation. The blind channel estimation needs to acquire a large amount of signal information, which has high computational complexity and poor real-time performance, and cannot meet the requirements of 5G users with low latency. In semi-blind estimation, the initial values of channel parameters are obtained by a small number of pilots, and the likelihood estimation of complete channel parameters is carried out by signal feedback combined with the posterior information. The method of channel estimation based on pilot frequency is to insert training sequence or pilot symbol which is not carrying useful information and is known to both sender and receiver in the process of signal transmission, and the receiver realizes channel estimation according to pilot frequency signal. Compared with semi-blind estimation and blind estimation, pilot-based signal estimation has the advantages of low complexity and simple method, so it is widely used.

Since the effectiveness of test-based channel evaluation depends on the design and calculation algorithm, many researchers have done a lot of research on the above two factors. Fernandes, P. B. For uncorrelated Ruyli fading channel, the optimal signal must be orthogonal and the test length must be equal to the antenna, which is difficult to use in 5G array multiple-input-multiple-output (MIMO) systems [5]. Zheng, R. Z. Channel feedback is used to complete the signal adjustment, which should reduce the interference between the interference signal and the interference signal generated by the receiver, thereby improving the accuracy of channel estimation [20]. Yukun uses the physical connection and proximity of the channel and uses the correlation method to reduce the mean square error of the channel estimation and optimize the test sequence [15]. Deng, X. Adopted the method of channel second-order statistics to allocate pilot frequency, so as to ensure that users using the same pilot frequency correspond to mutually orthogonal channel matrix, thus avoiding the problem of channel pollution [4]. Wang and S. proposed the channel state information feedback scheme of deep neural network based on the channel time correlation, which effectively reduced the channel state information feedback overhead and improved the channel state information feedback accuracy [14]. Sun and W. Z. used sparse channel estimation based on compressed sensing to obtain better channel estimation results, which could better resist pilot interference [11]. Chinnadurai, G. proposed a channel estimation method based on sparse conversion method, which could realize multi-channel parameter estimation under the condition of limited pilot overhead [3].

PSO algorithms were originally developed to simulate bird image quality and unknown motion. From the analysis of animal behavior, the correlation between the information of the group gives the results of the change, which is the basis for the development of the algorithm. The first threshold of PSO was calculated by adding the speed to the neighbors, taking into account the multi-detection and distance speed. Then inertia weights  $w$  are introduced to better control effort and search, so the model can be built.

The above pilot design and estimation algorithms are all designed based on small-scale MIMO-OFDM or SISO-OFDM channel system, the dynamic changes of the environment, system overhead and computational complexity in massive multiple antenna systems are not considered, so it cannot be directly extended to massive MIMO systems. Aiming at this problem, the author proposes a channel estimation method for 5G communication system based on improved particle swarm optimization algorithm, in order to reduce the system overhead and computational complexity of channel estimation, the channel estimation model is defined as regression fitting [16]. In order to deal with the dynamic changes of the environment and the inconsistency between the actual scene and the training data, the model is iteratively optimized by the method of online training of small sample data and integrated learning, so as to achieve high applicability of the model. The construction of a wireless communication network consists of three parts: ground stations, transmission stations, and transmission stations that can transmit communications and data around the world. answered. In wireless communication design, channel parity is important to ensure good wireless communication and communication quality. In strong interference and strong interference environment, communication is nonlinear, especially in hybrid MIMO mobile wireless communication, strong interference, it is difficult to know the channel register and channel balance of the communication system.

### 3. Methods.

**3.1. A standard particle swarm optimization algorithm.** Suppose  $I$  have  $m$  particles in a  $d$ -dimensional space, and each particle has a position vector and a velocity vector. Each location vector represents a solution. Its position and velocity at the  $k$ th iteration are  $X_i[k] = (x_{i1}, x_{i2} \dots x_{id})$  and  $V_i[k] = (v_{i1}, v_{i2} \dots v_{id})$ , respectively. The individual and global optimal values recorded by the whole particle swarm during the flight are  $p_{ibest}$  and  $g_{best}$ , respectively. Under the leadership of  $p_{itest}$  and  $g_{best}$ , the particle swarm optimizes the whole space. The basic formula of its evolution is (3.1) and (3.2):

$$V_i[k + 2] = wV_i[k] + c_1r_1(p_{iks} - X_i[k]) + c_2r_2(g_{best} - X_i[k]) \quad (3.1)$$

$$X_i[k + 2] = X_i[k] + V_i[k + 2] \quad (3.2)$$

where  $c_1$  and  $c_2$  are the individual and correlation coefficients, respectively, in general;  $r_1$  and  $r_2$  are two numbers equal to 0 and 1.  $w$  is the inertial weight used to measure the search and inspection ability of a small object. According to research, the inertia of a weight is often taken as a downward slope. Formula (3.3)

$$w = w_{\max} - (w_{\max} - w_{\min}) \frac{FEs}{T} \quad (3.3)$$

where  $w_{\max}$  and  $w_{\min}$  are the initial inertia weight values respectively;  $FEs$  is the iteration number of the function;  $T$  is the maximum number of iterations with linearly decreasing inertia weight;  $w$  remains the same after the number of iterations exceeds  $T$ . Usually  $w_{\max} = 0.9, w_{\min} = 0.4$ .

**3.2. Improved particle swarm optimization algorithm.** In the particle swarm optimization algorithm, the individual and the correlation coefficients  $c_1$  and  $c_2$  determine the global and local surveys of the population, respectively, and the following formula (3.4) can be obtained from the deformation of the equation (3.1) according to the data:

$$\begin{aligned} V[k + 1] &= wV[k] + c \{r_1(p_{thest} - X[k]) + r_2(g_{best} - X[k])\} \\ &= wV[k] + c(r_1 + r_2) \left\{ \frac{r_1}{r_1 + r_2} p_{iteses} + \frac{r_2}{r_1 + r_2} g_{best} - X[k] \right\} \\ &= wV[k] + ch \{h_1 p_{ikes} + (1 - h_1) g_{best} - X[k]\} \end{aligned} \quad (3.4)$$

where  $h = r_1 + r_2; h_1 = r_1 / (r_1 + r_2)$ , and the value range of  $h$  and  $h_1$  are  $[0, 2]$  and  $[0, 1]$  respectively.

If  $L[k] = h_1 p_{ikest} [k] + (1 - h_1) g_{hest} [k]$  is set, equation (3.5) is:

$$V[k + 1] = wV[k] + ch(L[k] - X[k]) \quad (3.5)$$

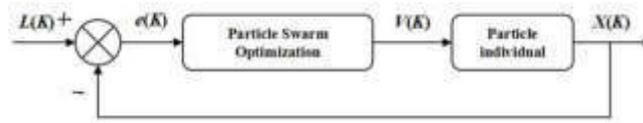


Fig. 3.1: Closed-loop control of Particle swarm optimization algorithm

In the formula,  $L[k]$  is the search function of the particle swarm optimization algorithm, and according to the PID control theory, the particle swarm optimization algorithm can be considered as a closed-loop feedback control strategy in the search,  $L[k]$  can be regarded as the given value in PID control, while  $X[k]$  can be regarded as the feedback value of the system. Each particle can be regarded as the controlled object, namely, the deviation  $e(k) = L[k] - X[k]$  of the negative feedback control system, the algorithm obtains the velocity vector of the particle according to the deviation signal, the whole particle swarm flies to the optimal value under the leadership of a given position  $L[k]$  ( $L[k]$  itself contains both individual and global optimal factors). As shown in Figure 3.1:

Substituting equation (3.5) into equation (3.6), we can get:

$$X[k + 2] = X[k] + wV[k] + ch(L[k] - X[k]) \tag{3.6}$$

Take the expectation of both sides at the same time to obtain Equation (3.7)-(3.9):

$$EX[k + 1] = EX[k] + wEV[k] + c(L' - EX[k]) \tag{3.7}$$

Where,  $L' = (p_{ibest} + g_{hest}) / 2$ , (because  $Eh = 1, Eh_1 = \frac{1}{2}$ ), the following can be deduced:

$$EX[k + 1] - EX[1] = w(EX[k] - EX[0]) + c \sum_{j=0}^k (L' - EX[j]) \tag{3.8}$$

By transforming equation (3.8) according to the definition of deviation  $e(k)$ , we can obtain:

$$EX[k + 1] = EX[k] + (1 - w)Ee[k] + c \sum_{j=0}^k Ee[j] + X[1] - wX[0] - (1 - w)L' \tag{3.9}$$

That is:

$$EV[k + 1] = (1 - w)Ee[k] + c \sum_{j=0}^k Ee[j] + X[1] - wX[0] - (1 - w)L' \tag{3.10}$$

Equation (3.10) shows that it is similar to the discrete PI governing equation, where  $K_P = 1 - w, K_1 = c$ . Therefore, PSO can be regarded as a PI controller,  $V[k]$  as the output of the controller, and the dynamic characteristics of PSO can be described as Formula (3.11)-(3.13):

$$X[k + 1] - X[k] = \Delta X[k] = V[k + 1] \tag{3.11}$$

The expression of its discrete domain is as follows:

$$\frac{dX(t)}{dt} = V(t) \tag{3.12}$$

Its open-loop transfer function is:

$$G(s) = \frac{X(s)}{V(s)} = \frac{1}{s} \tag{3.13}$$

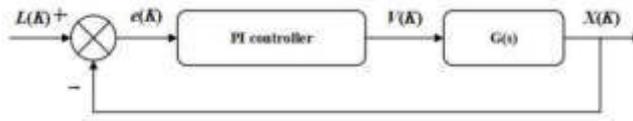


Fig. 3.2: Control block diagram of standard particle swarm optimization algorithm

A control block diagram of the standard particle swarm optimization algorithm is shown in Figure 3.2. The open-loop transfer function of PI controller is shown in equation (3.14):

$$G_{PI}(s) = K_P + K_I/s \quad (3.14)$$

The closed-loop transfer function is expressed in equation (3.15):

$$\varphi_{PI}(s) = \frac{G_{PI}(s)G(s)}{1 + G_{PI}(s)G(s)} = \frac{K_P s + K_I}{s^2 + K_P s + K_I} \quad (3.15)$$

Then, the following factors mainly affect the performance of the particle swarm optimization algorithm.

(1) Overshoot, which will affect the search ability of particle swarm optimization algorithm. Large overshoot means large particle search range and strong exploration ability.

(2) Attenuation ratio, a large attenuation ratio represents that the particle will be closer to the optimal value at a larger speed, the search range of the algorithm becomes smaller, but a small attenuation ratio will lead to the particle may be difficult to converge [19, 2, 17].

Usually, taking large scale coefficient and integral coefficient will make the algorithm have large overshoot and small decay ratio. Because the differential coefficient has the function of feedforward control, the overshoot and decay ratio can be reduced, and the differential effect should be reduced or removed in the later stage of the control for the stability of the system [8]. In this way, the convergence speed of the algorithm can be accelerated. The particles fly to the optimal position at a relatively fast speed in the beginning stage, and maintain a good global search ability in the later stage.

**3.2.1. Improving policies.** According to the above analysis, the particle swarm optimization algorithm is a basic PI controller, on the basis of which the difference between the parameters of the control parameters is shown in the first stage of the algorithm, the ratio of the coefficients and the balance. To improve the speed of rotation of the particle swarm optimization algorithm, small proportions and combinations are brought to the next stage, and the differences of the differences are reduced and the decision differences are shown according to their control. scientific world [13, 10]. The standard equation of its discrete domain is equation (3.16):

$$D[k] = K_D(e[k] - e[k - 1]) \quad (3.16)$$

The incremental differential action equation is equation (3.17):

$$\begin{aligned} \Delta D[k] &= D[k] - D[k - 3] = K_D(e[k] - e[k - 1]) - K_D(e[k - 2] - e[k - 2]) \\ &= K_D\{(L[k] - x[k]) - 2(L[k - 1] - X[k - 1]) + (L[k - 2] - X[k - 1])\} \end{aligned} \quad (3.17)$$

The expression of standard particle swarm optimization algorithm with incremental differential function is equation (3.18):

$$V_i[k + 1] = wV_i[k] + c_1 r_1 (p_{ib \text{ best}} - X_i[k]) + c_2 r_2 (g_{\text{best}} - X_i[k]) + \Delta D_i \quad (3.18)$$

Therefore, the iterative process of improving the particle swarm optimization algorithm is Equation (3.19):

$$\begin{aligned} V_i[k + 1] &= wV_i[k] + c_1 r_1 (p_{ibest} - X_i[k]) + c_2 r_2 (g_{\text{best}} - X_i[k]) + K_D\{(L[k] - L[k - 1]) \\ &\quad - (L[k - 2] - L[k - 3]) - (V[k] - V[k - 1])\} \end{aligned} \quad (3.19)$$

where  $L[k] = (p_{\text{itest}} + g_{\text{test}}) / 2$ .

**3.2.2. Convergence analysis of improved particle swarm optimization algorithm.** Based on equations (3.2), (3.17) and (3.20), we can obtain:

$$\begin{aligned} c_1 r_1 p_{\text{bbest}} + c_2 r_2 g_{\text{best}} &= (c_1 r_1 + c_2 r_2 - 1 - w + K_D) X_i[k+2] + X_i[k+3] \\ &+ (w - 2K_D) X_i[k+1] + K_D X_i[k] \end{aligned} \quad (3.20)$$

Assuming that the individual and global optimal values of particles remain unchanged, equation (3.21)-(3.23) can be obtained by taking both sides of the equation as expected:

$$\begin{aligned} (c_1 p_{\text{ikest}} + c_2 g_{\text{hest}}) / 2 &= \left( \frac{c_1 + c_2}{2} - 1 - w + K_D \right) EX_i[k+2] + EX_i[k+3] \\ &+ (w - 2K_D) EX_i[k+1] + K_D EX_i[k] \end{aligned} \quad (3.21)$$

After  $z$  transformation:

$$(z^3 + T_1 z^2 + T_2 z + T_3) EX(z) = S \quad (3.22)$$

where  $T_1 = (c_1 + c_2) / 2 - 1 - w + K_D$ ;  $T_2 = w - 2K_D$ ;  $T_3 = K_D$ ;  $S = \frac{c_1 p_{\text{ibest}} + c_2 g_{\text{best}}}{2} + z^3 X(0) + z^2 [X(1) + T_1 X(0)] + z [X(2) + T_1 X(1) + T_2 X(0)]$ .

$z = (u + 1) / (u - 1)$  is exchanged by linear difference variation, so:

$$Au^3 + Bu^2 + Cu + D = S \quad (3.23)$$

Type,  $A = 1 + T_1 + T_2 + T_3$ ;  $B = 3 + T_1 - T_2 - 3T_3$ ;  $C = 3 - T_1 - T_2 + 3T_3$ ;  $D = 1 - T_1 + T_2 - T_3$ . The characteristic equation of equation (3.24) is:

$$Au^3 + Bu^2 + Cu + D = 0 \quad (3.24)$$

As for stability conditions, a system is stable only if the following conditions are satisfied.

$$\begin{cases} A > 0 \\ B > 0 \\ \frac{BC-AD}{B} > 0 \\ D > 0 \end{cases} \Rightarrow \begin{cases} c_1 + c_2 > 0 \\ 4 - 4w + (c_1 + c_2) > 0 \\ 2(1 - w)(1 + K_D) + (c_1 + c_2)K_D > 0 \\ 4 + 4w - (c_1 + c_2) - 8K_D > 0 \end{cases} \quad (3.25)$$

Under the condition that equation (3.25) is satisfied, the algorithm converges to  $(c_1 p_{\text{ithest}} + c_2 g_{\text{best}}) / (c_1 + c_2)$ , namely equation (3.26):

$$EX(\infty) = \lim_{z \rightarrow 1} (z - 1) EX(z) = \frac{c_1 p_{\text{ibest}} + c_2 g_{\text{best}}}{c_1 + c_2} \quad (3.26)$$

From the above analysis, it is known that the improved particle swarm optimization algorithm converges under the premise of satisfying equation (3.26).

**3.3. Adaptive optimization model for channel estimation.** Particle adaptive control and convolution measure extraction of communication signals based on signal and channel analysis, channel equalization is improved. The initial channel estimation model is constructed by training the channel state matrix from a large number of offline data, which is difficult to be used directly in the dynamic channel time-domain environment [9]. In addition, there is a certain relationship between the channel estimation model and the service volume of users, at different time scales (busy and idle), the external interference of channel estimation is also different, therefore, the initial channel estimation model based on offline big data cannot be used for channel estimation in different regions and different time periods, otherwise, its prediction accuracy and performance will be greatly compromised. To solve this problem, the authors optimized the first channel prediction by online training. Since the transmission and reception data are constantly updated, a new channel prediction model is created by training new data, and the weight of the original channel prediction model is changed by comparing the performance of the original channel prediction models. An updated weighted average method is used to adjust the weight of each forecast channel model, and an online performance evaluation model is developed to learn the learning curve of the forecast channel [1]. Those. Figure 3.3 shows the channel prediction optimization model based on online learning.

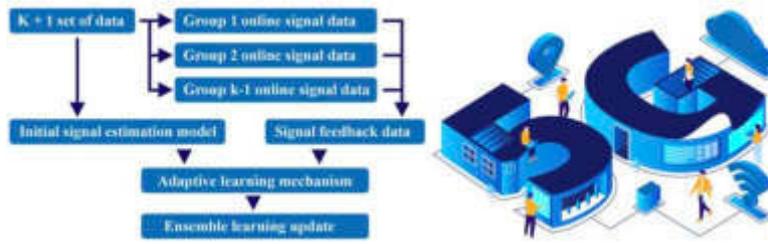


Fig. 3.3: Channel estimation optimization model based on online ensemble learning

Table 3.1: Related channel parameters in simulation experiments

Parameter	Value
Number of subcarriers $L$	6
The antenna number $M$	100
Number of users per cell $k$	15 ~ 20
Channel propagation path $P$	100
Carrier frequency $f/\text{GHz}$	2.5
The wavelength $\lambda/m$	0.12
The antenna spacing $d/m$	0.06
The antenna impedance / $\Omega$	50
Transmission of elevation	[-90,90]
Distance between the user and the base station $l_1/m$	100
Protection distance between user and base station $l_2/m$	10
Path loss correlation coefficient	3.8
Number of pilot	10 ~ 100
Transmission power of the user terminal $P/\text{dBm}$	-10

**3.4. Simulation experiment.** In order to verify the author's model, two classical channel estimation methods, LS channel estimation and BLMMSE channel estimation, are compared with the author's model, the channel estimation accuracy of the two scenarios with different SNR and different pilot lengths will be compared. LS channel estimation realizes channel estimation by least square method. BLMMSE channel estimation is to decompose the signal into linear variables, and then realize the channel estimation by least square method. It can be seen that BLMMSE channel estimation can better deal with the problem of nonlinear mapping and has high estimation accuracy.

The relevant channel parameters used in simulation are shown in Table 3.1.

**4. Results and Discussion.** Comparison of estimation accuracy of three estimation methods under different SNR, Figure 4.1 shows that the main coordinate axis corresponds to the errors of LS channel estimation and BLMMSE channel estimation. The subcoordinate axis corresponds to the channel estimation error. The performance of the channel estimation model is optimal under different channel ratios, with the increase of SNR, the MSE value is stable at about 1200. LS channel estimation and BLMMSE channel estimation methods are not only sensitive to SNR changes, but also have low estimation accuracy at low SNR. This is because in the state of low SNR, the label values corresponding to different channels have little difference, it is difficult to establish an accurate mapping relationship between LS channel estimation and BLMMSE channel estimation when the label difference is small. The authors' model uses online data to continuously optimize the model, an accurate mapping relationship is constructed by iterative method to realize adaptive learning of channel estimation parameters in the scenario of SNR fluctuation [7].

Comparison of the estimation accuracy of the three estimation methods under different pilot lengths, Figure 4.2 shows that the main coordinate axis corresponds to the error of LS channel estimation and BLMMSE channel estimation. The sub-coordinate axis corresponds to the author's channel estimation error. The performance

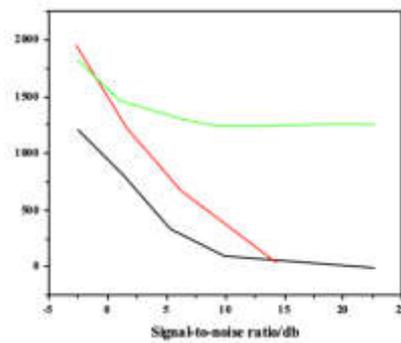


Fig. 4.1: Comparison of estimation accuracy of the three estimation methods under different SNR

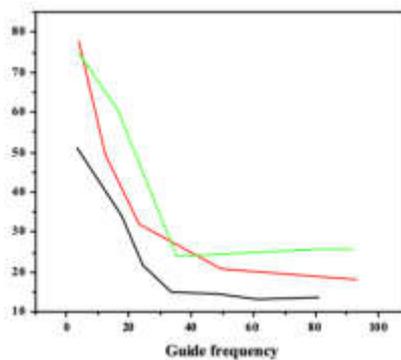


Fig. 4.2: Comparison of estimation accuracy of the three estimation methods under different pilot lengths

of the author’s channel estimation model is optimal under different pilot lengths, the LS channel estimation and BLMMSE channel estimation methods are not only relatively sensitive to the change of pilot length, but also when the pilot length, due to the high number of antennas (the number of antennas is 100), its channel estimation accuracy is low. The author divides every five sampling gaps into a data block, when the first data block under each sampling gap is designed as pilot, the other four channel data are estimated by the first data block, in the case of more antennas and low pilot length, the smaller pilot length has little difference, in the case of small tag difference, it is difficult to establish accurate mapping relationship between LS channel estimation and BLMMSE channel estimation [12, 18]. The author’s model uses online data to continuously optimize the model and builds an accurate mapping relationship through iteration, which can better cope with the changes of pilot length.

**5. Conclusion.** In the 5G communication model, channel balance is the key to ensure wireless communication quality and smooth communication. In the environment of strong interference and interference, the communication system has unique characteristics, especially the hybrid MI-MO mobile wireless communication system is prone to strong interference, and thus it is difficult to register the communication channel and achieve channel balance. In this paper, a flame-based fitness monitoring is proposed to test the robustness and performance of the algorithm in terms of signal-to-noise ratio and length difference.

The author introduces channel estimation for 5G communication based on network optimization and divides the channel estimation problem into two problems: regression optimization and online adaptive ensemble optimization. The improved particle swarm optimization algorithm is used to train a lot of data offline, and the first model of the wireless channel is obtained by learning the characteristics of each state of the wireless channel. Based on this, the model was updated with online data, and the modification of the forecast channel was done

by the joint method. Finally, the authors create experiments to verify the effectiveness and performance of the algorithm with different SNRs and different lengths. The results show that the method proposed in this paper is the most effective in balancing 5G communication. The error is small, the balance of error is small, the impact is small, and the accuracy of communication has been improved. In conclusion, the channel evaluation method proposed by the authors is feasible and effective.

## REFERENCES

- [1] F. S. ABDULLA, A. N. HAMOODI, AND A. M. KHEDER, *Particle swarm optimization algorithm for solar pv system under partial shading.*, *Przeglad Elektrotechniczny*, 97 (2021).
- [2] M. ASADI, M. A. J. JAMALI, S. PARSA, AND V. MAJIDNEZHAD, *Detecting botnet by using particle swarm optimization algorithm based on voting system*, *Future Generation Computer Systems*, 107 (2020), pp. 95–111.
- [3] G. CHINNADURAI AND H. RANGANATHAN, *Curvature based 2 wheels self supporting robot based on the particle swarm optimization algorithm*, *International Journal of Computing & Information Technology*, 11 (2019), pp. 37–43.
- [4] X. DENG AND Y. LIN, *Improved particle swarm optimization for mean-variance-yager entropy-social responsibility portfolio with complex reality constraints*, *Engineering Computations*, 39 (2022), pp. 1288–1316.
- [5] P. B. FERNANDES, R. OLIVEIRA, AND J. F. NETO, *Trajectory planning of autonomous mobile robots applying a particle swarm optimization algorithm with peaks of diversity*, *Applied Soft Computing*, 116 (2022), p. 108108.
- [6] R. HUANG, J. FAN, AND H. TSAI, *Application of dormancy particle swarm optimization algorithm based on nonlinear adjustment parameters for maximum power point tracking of photovoltaic array*, *International Journal of Electrical Engineering: Transactions of the Chinese Institute of Engineers*, 28 (2021), pp. 149–159.
- [7] Z. LIN, X. FU, B. GU, AND Z. FU, *Retracted article: Groundwater pollution prevention based on improved particle swarm algorithm and sports training optimization*, *Arabian Journal of Geosciences*, 14 (2021), pp. 1–13.
- [8] K. SAPLIOGLU, T. S. K. OZTURK, AND R. ACAR, *Optimization of open channels using particle swarm optimization algorithm*, *Journal of Intelligent & Fuzzy Systems*, 39 (2020), pp. 399–405.
- [9] J. SHOBANA AND M. MURALI, *Adaptive particle swarm optimization algorithm based long short-term memory networks for sentiment analysis*, *Journal of Intelligent & Fuzzy Systems*, 40 (2021), pp. 10703–10719.
- [10] J. SUBUDHI AND P. INDUMATHI, *Joint user clustering and salp based particle swarm optimization algorithm for power allocation in mimo-noma*, *Journal of Intelligent & Fuzzy Systems*, 40 (2021), pp. 9007–9019.
- [11] W.-Z. SUN, M. ZHANG, J.-S. WANG, S.-S. GUO, M. WANG, AND W.-K. HAO, *Binary particle swarm optimization algorithm based on z-shaped probability transfer function to solve 0-1 knapsack problem*, *IAENG Int. J. Comput. Sci*, 48 (2021).
- [12] M. SUNDARAY, A. TRIPATHY, AND S. TRIPATHY, *A new algorithm based on particle swarm optimization for application in holographic coupler*, *Optik*, 208 (2020), p. 164551.
- [13] B. URUL, *High-gain uwb antenna with optimized metamaterial lens layer via particle swarm optimization algorithm for satellite communication applications*, *International Journal of RF and Microwave Computer-Aided Engineering*, 32 (2022), p. e23280.
- [14] S. WANG, P. LI, H. JI, Y. ZHAN, AND H. LI, *Prediction of air particulate matter in beijing, china, based on the improved particle swarm optimization algorithm and long short-term memory neural network*, *Journal of Intelligent & Fuzzy Systems*, 41 (2021), pp. 1869–1885.
- [15] Y. WANG AND X. CHEN, *Hybrid quantum particle swarm optimization algorithm and its application*, *Science China Information Sciences*, 63 (2020), pp. 1–3.
- [16] S. G. WARDHANA AND W. PRANOWO, *Rock-physics modeling by using particle swarm optimization algorithm*, *Journal of Applied Geophysics*, 202 (2022), p. 104683.
- [17] C. YONG AND Y. LU, *An improved particle swarm optimization algorithm in selection of e-commerce distribution center*, *Journal of Intelligent & Fuzzy Systems*, 39 (2020), pp. 8783–8793.
- [18] L. ZHANG, J. WANG, AND Z. AN, *Fcm fuzzy clustering image segmentation algorithm based on fractional particle swarm optimization*, *Journal of Intelligent & Fuzzy Systems*, 38 (2020), pp. 3575–3584.
- [19] Q. ZHANG, Y. WEI, W. SONG, ET AL., *Two strategy cooperative particle swarm optimization algorithm with independent parameter adjustment and its application*, *International Journal of Innovative Computing, Information and Control*, 16 (2020), pp. 1203–1223.
- [20] R. ZHENG, Y. ZHANG, AND K. YANG, *A transfer learning-based particle swarm optimization algorithm for travelling salesman problem*, *Journal of Computational Design and Engineering*, 9 (2022), pp. 933–948.

*Edited by:* B. Nagaraj M.E.

*Special issue on:* Deep Learning-Based Advanced Research Trends in Scalable Computing

*Received:* Aug 14, 2023

*Accepted:* Oct 17, 2023



## MULTI-SOURCE AND MULTI-LEVEL COORDINATION OF ENERGY INTERNET UNDER V2G BASED ON PARTICLE SWARM OPTIMIZATION ALGORITHM

JIAN XU\*, YUNYAN CHANG<sup>†</sup> AND XIAOMING SUN<sup>‡</sup>

**Abstract.** In order to effectively improve the excessive load of microgrid during peak hours of urban electricity consumption, a multi-source and multi-level coordination method of energy Internet under V2G based on particle swarm optimization algorithm was proposed. First, a mathematical model for V2G energy integration in microgrids was developed and a scheduling concept based on a particle-by-particle optimization algorithm was used. Second, an improved PSO algorithm is proposed and experimentally validated, and the experimental results are compared with previous particle swarm optimization algorithms. Experiments have shown that as the number of iterations increases, the value of the objective function decreases and the optimal solution can be obtained until the maximum number of iterations is reached. The iteration speed and power processing cost of the improved PSO algorithm are better than before. The original load curve is the load trough period from 23:00 to 6:00, and two load peaks occur from 12:00 to 14:00 and 19:00 to 22:00. The V2G technology basically realizes the coordinated control of microgrid electric energy and achieves the effect of peaking and valley filling. The improved algorithm has obvious improvement compared with the original power grid state. Conclusion: The application of EV V2G technology can smooth the daily load curve of power grid and coordinate the electric energy of micro-grid to achieve “peak cutting and valley filling”, and the effect of this algorithm is more outstanding than the previous algorithm. Finally, the future development direction and suggestions of V2G technology are put forward. The power grid with V2G discharge depth limit has the ability to basically reduce and eliminate the daily peak load, so the technology has broad research space and development prospects.

**Key words:** V2G technology, Microgrid energy control, Power network peak regulation, Load curve, Particle swarm optimization

**1. Introduction.** At present, due to the rapid development of the global economy and the increasingly prominent population problem, energy shortage has become an urgent problem to be solved. With the increasingly serious ecological environment problems, saving energy, reducing pollution emission and developing low carbon economy have become an important task of the development of human society. With the diversified application of energy more digital and intelligent, people connect billions of facilities, transmission and consumption of energy production, systems and information through advanced sensors, control and software applications, forming the energy Internet. China is the second largest oil consumer in the world. At a time when there is a shortage of fuel and increasing environmental pollution, the country supports construction to save money and protect the environment. According to the statistics of the Environmental protection department, the carbon dioxide content of automobile exhaust emissions is second only to the carbon emissions of traditional energy processing and energy reuse, accounting for about a quarter of the global total carbon emissions [12]. In the automobile industry, the exhaust emission of traditional fuel vehicles is one of the most important pollution sources. According to the auto sales department, China’s auto production and sales have been growing at an annual rate of 1 million since 2000, and the number of cars in China is expected to reach at least 550 million by the middle of the 21st century, which will be at least 38% higher than the number of cars in the United States during the same period. The rapid development of automobile industry not only brings convenience and business opportunities to people, but also aggravates energy shortage and environmental deterioration. In the past, the state has promoted energy conservation, emission reduction, low-carbon economy development, and low-carbon life, and many researchers have dedicated to the development of low-carbon transportation [8]. Those. China’s fifth-year plan emphasized the “development of green and low-carbon construction concepts” and included new energy vehicles in the “seven emerging economic strategies”. China’s new energy industry has

---

\*Chongqing Water Resources And Electric Engineering College, Chongqing, China (Corresponding author, [JianXu37@163.com](mailto:JianXu37@163.com))

<sup>†</sup>Chongqing Water Resources And Electric Engineering College, Chongqing, China ([YunyanChang2@126.com](mailto:YunyanChang2@126.com))

<sup>‡</sup>Chongqing Water Resources And Electric Engineering College, Chongqing, China ([XiaomingSun85@163.com](mailto:XiaomingSun85@163.com))

gradually entered the stage of economic development. With the continuous development of battery technology, compared with fuel vehicles with high pollution and high energy consumption, electric vehicles have attracted more and more attention due to their advantages of energy saving, environmental protection and low noise. Their large-scale application is regarded as an effective way to alleviate energy shortage and air pollution and promote the development of low-carbon economy. Electric vehicles will also be an inevitable trend for the development of the automobile industry in the future [19, 7]. Due to the different priorities considered, the objective functions and constraints of energy Internet scheduling optimization will take various forms, but they can be attributed to non-linear multi-objective planning problems, which can be solved by optimization algorithm, such as genetic algorithm, immune algorithm, particle swarm algorithm, etc.

In recent years, countries around the world have carried out electric vehicle development plans. For example, the United States plans to produce more than 1 million electric vehicles in 2015, Germany plans to increase its domestic electric vehicle ownership to 1 million in 2020, and Denmark plans to gradually replace traditional fuel vehicles with electric vehicles [15, 11]. With the rapid development of new energy generation technologies such as renewable energy, energy efficient and clean fossil fuels, micro-distributed power generation is a gradual way to meet the growing demand for electricity, reduce greenhouse gas emissions and improve energy efficiency. Domestic research institutions also began to conduct relevant investigation on electric vehicles. As early as 1995, the Ministry of Science and Technology of China carried out the electric vehicle project of major science and technology industry. Hybrid and pure electric vehicles were already available in China in early 2001. From 2001 to 2005, the national 863 Program set up various special research and development funds for electric vehicles reached 3 billion yuan. The electric vehicle application project launched in Beijing Olympic Games has played a good role in driving and demonstrating the promotion of electric vehicles. In 2009, China launched a demonstration project of "1,000 vehicles in ten cities", involving 13 cities. First, new energy vehicles, mainly plug-in hybrid vehicles, were promoted and used in public transport. In 2010, 25 cities participated in the demonstration project of "Thousand Cars in Ten cities", and private electric vehicles were promoted in 6 cities [13]. Compared with traditional algorithms, the Particle Swarm Optimization (PSO) algorithm has advantages such as simplicity, ease of use, fast rotation, and global optimization.

**2. Literature Review.** With the advancement of science and technology, the share of electric vehicles will increase. In the future, many electric vehicles will be connected to the electric power through the distribution of medium and small. These electric vehicles account for most of the total load on the grid, which will have a significant impact on the quality of energy and the safety and stability of the financial grid. Evidence shows that most electric vehicles do not work for most of the day, and using the idle time of the electric batteries will help to make the most of the project. With the rapid development of electric vehicle technology, the number of electric vehicles continues to increase. If many electric vehicles with poor charging ability enter the grid, it will seriously affect the safe and stable operation of the electric grid [3]. This problem worries domestic and foreign scientists. Many countries have studied the management and control of large-scale electric vehicles entering the grid. According to statistics, most electric cars are closed for 90 percent of the day, and the energy storage of electric cars has been decided. The emergence of V2G technology brings solutions to the above problems. The main purpose of V2G is to enable two-way communication between data and power between electric vehicles and the grid. Electric vehicles can leave the grid at higher speeds during peak load times and charge less during low times, so they can even out the system, cutting peaks and filling valleys. Those. Electric vehicle users can find a difference in energy costs, and the grid can save operating costs by reducing resources. Continued research on V2G. Also, smartly connect the electric car to the grid to get the dispatch instructions. Since then, electric vehicles have become part of grid-friendly transportation. The results of the study of electric vehicles participating in the electric frequency network were established abroad. The participation of electric vehicles in the frequency adjustment of the electric power grid is often related to feasibility, economic and management strategies. Although the concepts are different, their main purpose is to provide theoretical support for electric vehicles for high-frequency services [6, 5]. The Particle Swarm Optimization Algorithm is an adaptive technology based on swarm intelligence inspired by artificial life research. The basic understanding of the history of particle flocking comes from studies of the predatory behavior of bird flocks. Researchers have found that birds often suddenly change direction, explosions, their behavior should not be expected, but in all situations, the maximum distance between people. By studying the behavior of similar groups, he discovered

that shared information about relationships exists in biological groups, which is advantageous for changing group clusters, which is the basis for building particle swarm algorithms. In this paper, a mathematical model of V2G cooperation in advanced coordination is developed. Based on the planning strategy, a particle-by-particle optimization algorithm is proposed. The experimental results of the two algorithms were compared with MATLAB/Simulink simulation software, and the ability to control the electric vehicle charging and microgrid power output control was analyzed. Experiments show that the value of the objective function decreases with the number of iterations, and the optimal solution is reached when the number of iterations reaches a maximum. The improved PSO algorithm is faster and more efficient. This shows that the particle swarm optimization algorithm is more efficient [2].

### 3. Research Methods.

**3.1. Overview of V2G technology.** The main purpose of V2G is to realize the two-way interaction of information and energy between electric vehicles and the power grid. Electric vehicles can discharge to the grid at a higher price when the peak load of the grid is high, and charge at a lower price when the grid trough, so as to smooth the system load and peak filling. Electric vehicle users can also earn the price difference, and the power grid also saves operating costs by reducing the reserve capacity. V2G technology fully embodies the friendly interaction between electric vehicles and smart grid, which is a two-way interaction technology of energy and information. There are many types of electric vehicles and they are very different in nature. Gasoline-electric hybrid electric vehicles have no devices to interact with the grid, so connecting to the grid will have no effect. Only pure electric vehicles and plug-in hybrid electric vehicles interact with the power grid, so V2G technology can be considered. The electric vehicles studied in this essay refer to these two types [1]. Since electric vehicles are idle in most of the day, if the V2G function of electric vehicles is utilized to make the energy flow bidirectional between the vehicle and the microgrid, the electric energy can be stored in the battery to meet its own demand when the load is used, and the grid can be used as a distributed energy storage device during peak load hours. V2G technology can not only reduce the network loss in the transmission process of long-distance high-volume electric energy, but also greatly reduce the thermal power capital needed to be put into the power grid at peak times. Taking advantage of all the changes in the smart grid and electric vehicle battery charging and charging, V2G technology can make the most adjustments to the power grid, smooth the daily load curve of the power grid, and reduce high voltage [4].

**3.2. 2V2G system components and information flow.** The V2G system model can be divided into four layers: network layer, parking management layer, smart charging and charging device layer, and vehicle layer. Among them, the bidirectional intelligent charging and discharging device not only interacts with the power grid, but also interacts with the vehicle. The operation of the whole system requires all devices to maintain communication. Figure 3.1 is the schematic diagram of information transmission [20].

**3.2.1. Battery management system.** Battery management system (BMS) is the main part of electric vehicle, including battery terminal module, central control module and display module. According to the different functions of each component, the BMS can store the power supply and current, the temperature, the state of health (SOC), the consumption state health (SOH), state of energy (SOE) and other pre-calculations and battery information. Fault self-check is carried out by diagnostic algorithm to prevent excessive charging and discharging from damaging vehicle battery, and timely alarm is given when necessary to ensure safe operation. BMS comprehensively monitors battery performance from all aspects, transmits data to charging and discharging devices (EV-PCS) with CAN bus, and then reports it to the background management system to make response strategies, and transmits charging and discharging instructions back to the charging and discharging devices, thus realizing the bidirectional flow of information [10, 16].

**3.2.2. Intelligent user terminal.** UT in Figure 3.1 is the user terminal, and users can intuitively understand the status information of electric vehicles on the dashboard. The component structure of the device is shown in Figure 3.2. The basic electrical equipment components include display, loudspeaker, card reader, etc. Most of the embedded ARM processors are used in the market, which mainly have wireless data communication module, GPS module and CAN bus interface. When the whole terminal system is installed on the EV, in addition to the bidirectional transmission of comprehensive battery information with the BMS mentioned

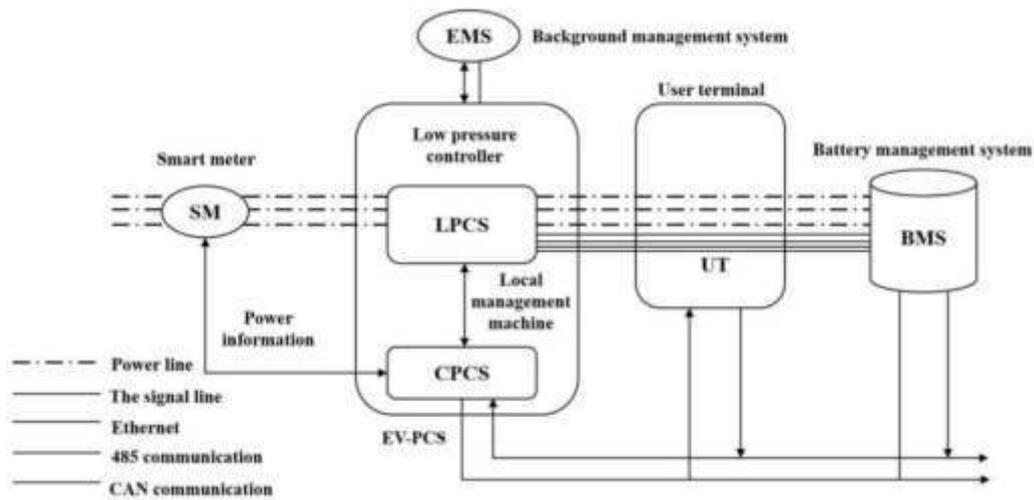


Fig. 3.1: V2G system information flow

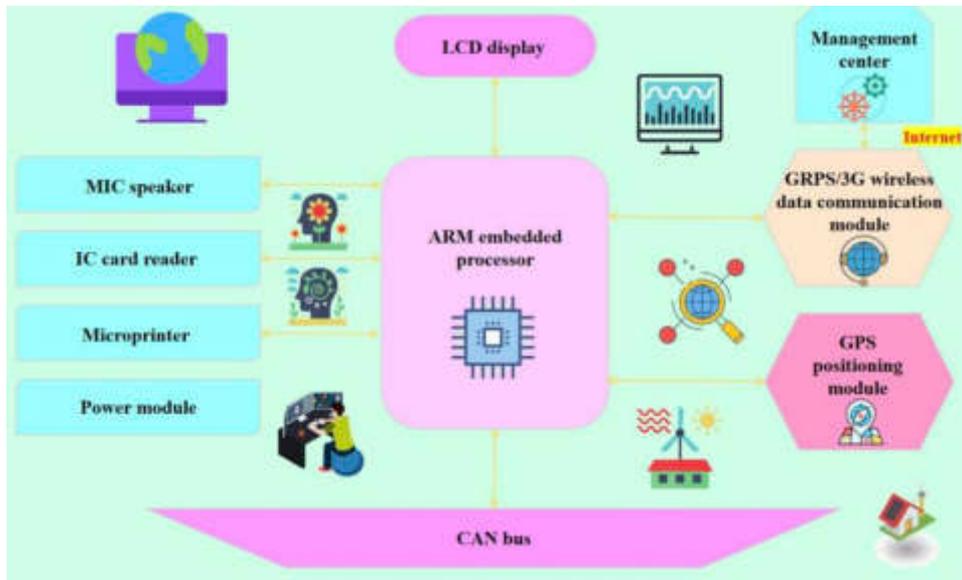


Fig. 3.2: User terminal structure

in Section 3.2.1, the user’s instructions can also be transmitted to the corresponding control devices, respectively realizing information interaction. Intelligent user terminal has GPS data reception function to locate the geographical location of the user; It has the wireless communication function and regularly transmits UT statistical information to the background management center. At present, it mainly uses GPRS/3G, a general wireless packet technology, and selects an appropriate network as the channel for remote communication with the outside world [17].

**3.2.3. Two-way intelligent charging and discharging device.** Intelligent charging and charging system EV-PCS consists of low power controller LPCS and local controller CPCS. At the core of V2G technology, EV-pcs communicate with all major devices. It is bidirectional connected to smart meter SM with RS485 to

record electricity information and set parameters. The bidirectional intelligent charging and discharging device receives the charging and discharging instructions from the control center wirelessly. In the charging mode, the execution object is only the vehicle, which is a one-way operation. In the case of V2G mode (peak load of the grid), the SOC set by the user needs to be obtained from UT, which is functionally combined with EMS into a system responsible for determining the charge and discharge strategy, operates bidirectional with BMS, and comes with protection functions such as undervoltage, overvoltage and overcurrent [14].

**3.3. V2G peak regulation mode.** V2G participation in power grid peak regulation has faster response speed and higher comprehensive benefit than the traditional mode. At present, the application strategy of V2G technology is mainly studied by using simulation software to predict the power grid load curve, with the purpose of smoothing the curve after electric vehicles enter the network.

**3.3.1. Objective function.** According to the optimization objective of “peak trimming and valley filling”, the control unit of the dispatching center was adjusted to 1h (24 units in total), and the minimum mean square error of the daily load curve was established as the objective function, i.e., Equation (3.1) below:

$$\min F = \sum_{j=1}^{24} \left( P_{Lj} - \frac{\sum_{j=1}^{24} P_{Lj}}{24} - \sum_{i=1}^n P_{ij} \right)^2 \quad (3.1)$$

where  $P_{Lj}$  is the power load of the grid at time period  $j$ ;  $\frac{\sum_{j=1}^{24} P_{Lj}}{24}$  represents the average daily load power of the grid;  $n$  is the number of electric vehicles connected to the power grid;  $P_{ij}$  is the total load power of  $i$  electric vehicles participating in the grid peak regulation during period  $j$ , and the discharge value is positive while the charging value is negative.

### 3.3.2. Constraints.

#### 1) Power constraint

The charge and discharge power constraint of electric vehicles is expressed in expression (3.2). Generally, the transmission power of the charging line is not allowed to exceed 15 kW. Therefore, the value of  $P_{\max}$  in the equation is 15, and the current constraint is mainly considered under this condition. In Formula (3.3) and (3.4),  $I_{ic}$  is the charging current;  $I_{id}$  is discharge current;  $I_{iN}$ , depending on the battery type of different types of electric vehicles, the rating is determined.

$$-P_{\max} \leq P_{ij} \leq P_{\max} \quad (3.2)$$

$$0 \leq I_{ic} \leq 1/3I_{iN} \quad (3.3)$$

$$0 \leq I_{id} \leq 2I_{iN} \quad (3.4)$$

Based on the above formula, the power constraints of electric vehicle  $i$  at moment  $j$  are Equations (3.5), (3.6) and (3.7), where  $V_{ij}$  is the rated voltage of vehicle charging (single-phase 220 V, three-phase 380 V).

$$P_{ij \min} \leq P_{ij} \leq P_{ij \max} \quad (3.5)$$

$$P_{ij \max} = \min(15, V_{ij} \cdot 2I_{iN}) \quad (3.6)$$

$$P_{ij \min} = \max\left(-15, -V_{ij} \cdot \frac{I_{iN}}{3}\right) \quad (3.7)$$

#### 2) Battery capacity constraints

The user can set the constraint range according to the driving demand, and the following equation (3.8) can be satisfied when the electric vehicle leaves the grid:

$$SO_{cset} \leq SO_c \leq Soc_c \quad (3.8)$$

Where, the charged state of the battery is represented by  $SO_c$ , which is defined as the ratio of the remaining capacity of the battery to the capacity  $Q_{iN}$  when it is fully charged.  $SO_{cset}$  is the state of charge constrained by the user on the battery;  $Soc_c$  is the full load state.

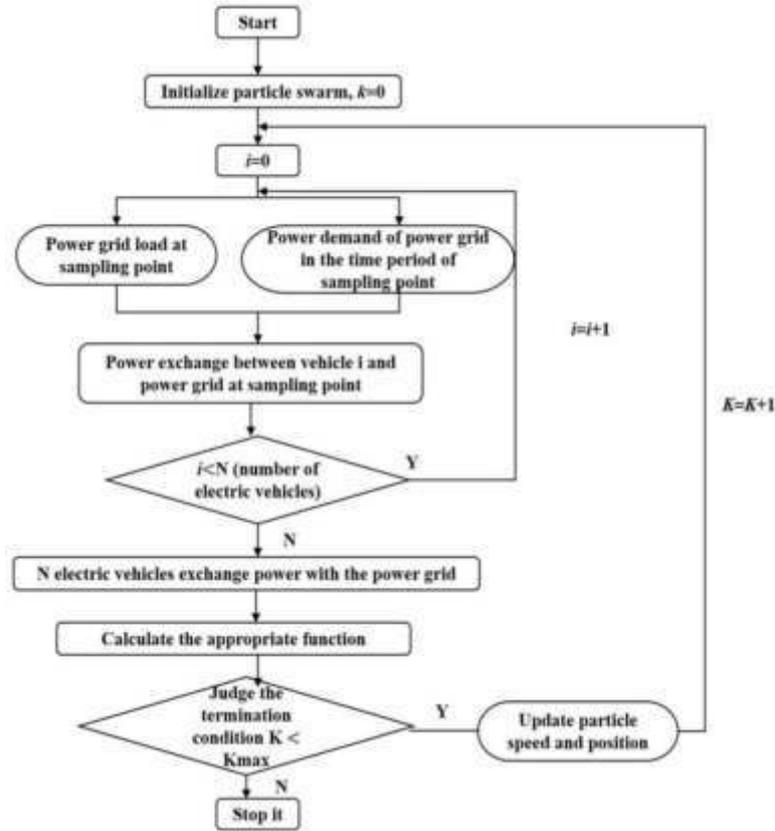


Fig. 3.3: Scheduling strategy flow based on particle swarm optimization algorithm

Considering the service life of the battery,  $Soc_c$  is required to be no less than 0.2 and no more than 1.  $S_{Socij}$  is the state of charge of vehicle  $i$  during the period  $j$ , and  $\Delta Q_{ij}$  is the change of battery capacity during charging and discharging.  $Q_{ij\max}$  represents the upper limit of capacity, and  $Q_{ij\min}$  represents the lower limit of capacity. Equation (3.9), (3.10), (3.11):

$$Q_{ij\min} \leq \Delta Q_{ij} \leq Q_{ij\max} \tag{3.9}$$

$$Q_{ij\max} = (S_{Socij} - S_{Socmin}) Q_{iN} \tag{3.10}$$

$$Q_{ij\min} = (S_{Socij} - S_{Socmax}) Q_{iN} \tag{3.11}$$

**3.4. Improved particle swarm optimization algorithm.** ParticleSwarmOptimization (PSO) or swarm foraging algorithm was first proposed by J. Kennedy and RK Eberhart. It is a modified version of a good genetic algorithm. The world’s best results have been achieved by implementing a random process, which has the advantages of fast rotation, high accuracy and ease of use. The different models developed in this letter are electric cars, and there are many uncertainties. The solution of the objective function is the minimum cost, and the particle-by-particle optimization algorithm can be used. As shown in Figure 3.3, the improvement in EV scheduling based on the particle swarm optimization algorithm is derived from the following two factors [9].

**3.4.1. Inertia weight of particles.** In the general search of the PSO algorithm, a larger inertia weight is suitable for the global search in the first stage of the iteration, while a small body inertia is useful for the local search in the later stage. To achieve a balance between the two, a line of weight reduction (LDIW) is

Table 4.1: Daily load of power network

Time period	Power /kW	Time period	Power /kW	Time period	Power /kW
1	66	9	78	17	83
2	66	10	90	18	97
3	64	11	88	19	103
4	67	12	102	20	92
5	68	13	102	21	103
6	67	14	97	22	93
7	74	15	83	23	71
8	79	16	80	24	68

proposed as shown in the following equation (3.12).

$$\omega = \omega_{\max} - (\omega_{\max} - \omega_{\min}) \frac{K}{K_{\max}} \quad (3.12)$$

where  $K$  is the number of iterations;  $K_{\max}$  is the set maximum number of iterations;  $\omega_{\max}$  is the maximum inertia weight;  $\omega_{\min}$  is the minimum inertia weight. With the increase of the number of iterations  $K$ , the inertia weight  $\omega$  gradually decreases.

The program added in Equation (3.12) can adaptively update the inertia coefficient. Select  $\omega_{\max} = 2$ ,  $\omega_{\min} = 0.1$ ,  $K_{\max} = 300$ .

**3.4.2. Particle renewal rate.** The situation of vehicles exchanging loads with the grid in 24 periods of a day is studied. Therefore, the dimension is 24, the number of initial population  $m = 100$ , and the position of particle  $i$  is expressed as  $X_i = (X_{i1}, X_{i2}, \dots, X_{i24})$ ,  $i = (1, 2, \dots, 100)$ . During operation, its "flight" speed is also  $n$ -dimensional vector  $V_i = (Vi1, Vi2, \dots, Vi24)$ ,  $i = (1, 2, \dots, 100)$ , and other speed parameters are set as  $V_{\min} = -0.5$ ,  $V_{\max} = 0.5$  according to experience.

Particle swarm optimization algorithm updates particle velocity according to equation (3.13), where  $x_i(k)$  is the position information of particle  $i$ ,  $p_i(k)$  represents individual extreme position of particle  $i$ , and  $g_i(k)$  represents global extreme position.

$$v_i(k+1) = \omega v_i(k) + c_1 r_1 (p_i(k) - x_i(k)) + c_2 r_2 (g_i(k) - x_i(k)) \quad (3.13)$$

In the speed update of the improved PSO, the speed of the last two times is weighted as the new speed. The main program is as follows:

$$\begin{aligned} [gbfi] &= \min(pbf) \\ gbx\_pso &= pbx(i, :) \\ Gbf\_pso &= [Gbf\_psogbf] \\ Gbx &= [Gbx; gbx\_pso] \\ v3(i, :) &= (1 - alf) * \left( \begin{array}{l} w * v2(i, :) + c1 * \text{rand} * (pbx(i, :) - P(i, :)) \\ + c2 * \text{rand} * (gbx\_pso - P(i, :)) \end{array} \right) + alf * v1(i, :) \end{aligned} \quad (3.14)$$

**4. Result Analysis.** The daily load curve of a building is used for simulation verification. The data are shown in Table 4.1, and  $P_{av}=83\text{kW}$  is obtained. For convenience of calculation, it is assumed that 2000 electric vehicles of the same model are connected to the power grid, and the driving power consumption is all 8kW. The rated capacity of the battery is  $40\text{kW} \cdot \text{h}$ , and the maximum allowable charging and discharging power is 15kW. The initial charging state of the electric vehicle entering the station is 0.8, and the minimum charging state of the outbound station required by the user is 0.9.

Table 4.2: Load exchange between electric vehicles and the grid

Time period	PSO	Improved PSO	Time period	PSO	Improved PSO	Time period	PSO	Improved PSO
1	3.9506	3.9967	9	7.4067	9.1229	17	8.9158	3.0122
2	6.7125	7.7567	10	-3.5161	-0.6706	18	-5.7265	-6.8803
3	7.6103	8.2171	11	-1.1951	-1.2442	19	-12.5054	-14.0272
4	9.6999	7.8026	12	-8.8.94	-11.5440	20	-6.0450	-0.3861
5	8.0242	7.9079	13	-8.1237	-11.4422	21	-5.8577	-11.3175
6	7.7998	8.3182	14	-9.1048	-3.6391	22	-2.4955	-1.4643
7	8.1648	7.6256	15	10.6419	9.3059	23	4.8398	18.4147
8	7.5815	7.2512	16	6.5271	9.0763	24	13.3846	19.6704

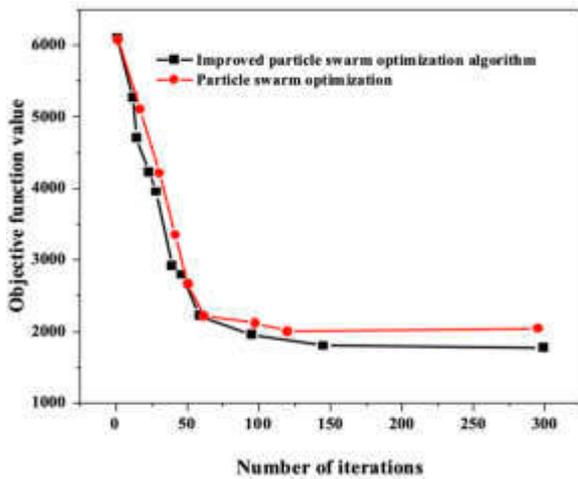


Fig. 4.1: Variation curve of objective function value with the number of iterations

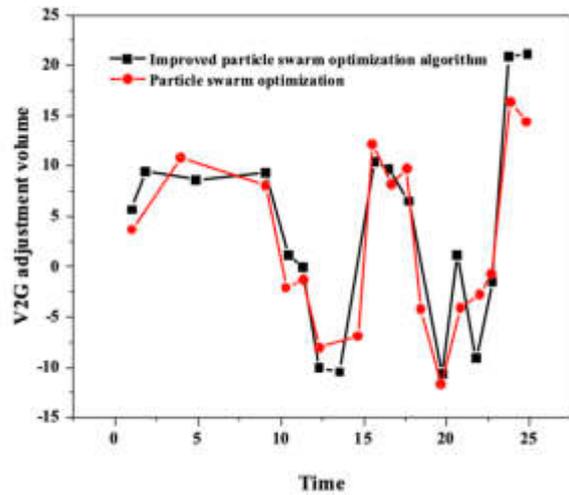


Fig. 4.2: V2G adjustment volume

According to the information in Table 4.1, based on the above mathematical model and the constraints, MATLAB simulation program is used to verify the accuracy of the results of the simple particle swarm optimization algorithm, and create the particle swarm optimization algorithm that is mentioned in this article. . follow. Those. The generated data is shown in Table 4.2 and the curves are shown in Figures 4.1 and 4.2.

Figure 4.1 shows that as the number of iterations increases, the value of the objective function decreases until the number of iterations reaches a maximum, and an agreement can be reached received. The iteration speed and fitness rate of the improved PSO algorithm are better than before. The values of the V2G rules in Figure 4.2 are put into the network, and the maximum rules of the two algorithms are compared as shown in Figure 4.3.

As can be seen from Figure 4.3, the load trough period of the original load curve is from 23:00 to 6:00, while two load peaks occur from 12:00 to 14:00 and 19:00 to 22:00. The simulation results show that V2G technology basically realizes the coordinated control of microgrid electric energy, and achieves the effect of peaking and valley filling. The improved algorithm is significantly improved compared with the original grid state, indicating that the algorithm is superior to the basic particle swarm optimization algorithm. The daily load curve can be replaced by a constant load curve, which is of great benefit to power distributors and reduces the price of electricity at the same time. It brings convenience to users [18].

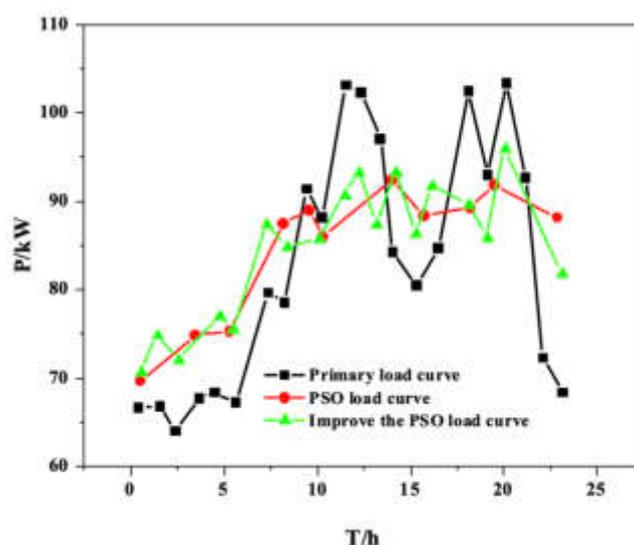


Fig. 4.3: V2G participating peak load regulation curve of power grid

**5. Conclusion.** In this essay, we fully demonstrate that the use of V2G can reduce the power grid load peak demand for generator sets, using the improved particle swarm optimization algorithm can achieve this point of view. The integration of electric vehicles into the power grid through V2G technology is not only beneficial to the distribution network in terms of power grid support, stability and load regulation, but also beneficial to the popularization of energy storage. However, improper V2G function management poses risks to the vehicle's built-in energy storage and requires setting discharge speed and depth to extend battery life. The improved PSO algorithm has better iteration speed and fitness function values. The original load curve is the load valley period from 23:00 to 6:00, 12:00~14:00 and 19:00~22:00. V2G technology basically realizes the coordinated control of microgrid power, and realizes the effect of peak value and filling valley. The improved algorithm has a significant improvement over the original grid state. A power grid with V2G discharge depth limit has the ability to reduce and eliminate the peak daily load essentially, so this technology has great research space and development prospect in the future. The existing research results show that V2G technology still has great potential for development, and the future research can focus on the specific implementation method of V2G technology, and the new business model suitable for V2G. V2G pilot projects can first be established in small-scale microgrids to expand their scale and extend them to the whole large power grid.

**Acknowledgement.** The study was supported by

1. Supported by the Natural Science Foundation of Chongqing Municipality (Grant No. cstc2021jcyj-msxmX1173).
2. Supported by the Science and Technology Research Program of Chongqing Municipal Education Commission (Grant No. KJZD-K202203801 KJQN202003804 KJQN202003807).
3. Supported by the Chongqing Educational Science Planning Project, Key topics, (Research on the integrated development mode of "industry city vocational innovation" in the western vocational education base) (Grant No. K22YC317058).
4. Supported by the Scientific Research Projects of Chongqing Water Resources And Electric Engineering College (Grant No. K202023, K202026).

#### REFERENCES

- [1] M. H. ALABDULLAH AND M. A. ABIDO, *Microgrid energy management using deep q-network reinforcement learning*, Alexandria Engineering Journal, 61 (2022), pp. 9069–9078.

- [2] M. AZEROUAL, Y. BOUJODAR, L. E. IYSAOUY, A. ALJARBOUH, M. FAYAZ, M. S. QURESHI, F. RABBI, AND H. E. MARKHI, *Energy management and control system for microgrid based wind-pv-battery using multi-agent systems*, Wind Engineering, 46 (2022), pp. 1247–1263.
- [3] H. GOLDSTEIN, *What v2g tells us about evs and the grid: Vehicle-to-grid technology adds another layer of complexity to the electric-vehicle transition*, IEEE Spectrum, 59 (2022), pp. 2–2.
- [4] R. JABEUR, Y. BOUJODAR, M. AZEROUAL, A. ALJARBOUH, AND N. OUAALINE, *Microgrid energy management system for smart home using multi-agent system*, International Journal of Electrical and Computer Engineering, 12 (2022), pp. 1153–1160.
- [5] M. KERMANI, P. CHEN, L. GÖRANSSON, AND M. BONGIORNO, *A comprehensive optimal energy control in interconnected microgrids through multiport converter under n- 1 criterion and demand response program*, Renewable Energy, 199 (2022), pp. 957–976.
- [6] M. S. MASTOI, S. ZHUANG, H. M. MUNIR, M. HARIS, M. HASSAN, M. ALQARNI, AND B. ALAMRI, *A study of charging-dispatch strategies and vehicle-to-grid technologies for electric vehicles in distribution networks*, Energy Reports, 9 (2023), pp. 1777–1806.
- [7] S. MOALEM, R. M. AHARI, AND G. SHAHGOLIAN, *Long-term demand forecasting in electrical energy supply chain of espidan ironstone industry using deep learning and extreme learning machine*, Journal of Intelligent Procedures in Electrical Technology, 13 (2022), pp. 1–20.
- [8] S. MOHSENI, R. KHALID, AND A. C. BRENT, *Metaheuristic-based isolated microgrid sizing and uncertainty quantification considering evs as shiftable loads*, Energy Reports, 8 (2022), pp. 11288–11308.
- [9] S. PANDA, S. MOHANTY, P. K. ROUT, AND B. K. SAHU, *A conceptual review on transformation of micro-grid to virtual power plant: Issues, modeling, solutions, and future prospects*, International Journal of Energy Research, 46 (2022), pp. 7021–7054.
- [10] M. S. RAFAQ, B. A. BASIT, S. A. Q. MOHAMMED, AND J.-W. JUNG, *A comprehensive state-of-the-art review of power conditioning systems for energy storage systems: Topology and control applications in power systems*, IET Renewable Power Generation, 16 (2022), pp. 1971–1991.
- [11] M. S. SADABADI, M. SHARIFZADEH, M. MEHRASA, H. KARIMI, AND K. AL-HADDAD, *Decoupled dq current control of grid-tied packed e-cell inverters in vehicle-to-grid technologies*, IEEE Transactions on Industrial Electronics, 70 (2022), pp. 1356–1366.
- [12] S. SAFIULLAH, A. RAHMAN, S. A. LONE, S. S. HUSSAIN, AND T. S. USTUN, *Robust frequency-voltage stabilization scheme for multi-area power systems incorporated with evs and renewable generations using ai based modified disturbance rejection controller*, Energy Reports, 8 (2022), pp. 12186–12202.
- [13] D. SONG, *A review of v2g technology applications and impact on the grid*, International Core Journal of Engineering, 8 (2022), pp. 640–645.
- [14] P. WANG, W. YUAN, C. SU, Y. WU, L. LU, D. YAN, AND Z. WU, *Short-term optimal scheduling of cascade hydropower plants shaving peak load for multiple power grids*, Renewable Energy, 184 (2022), pp. 68–79.
- [15] Z. WANG, P. JOCHEM, H. Ü. YILMAZ, AND L. XU, *Integrating vehicle-to-grid technology into energy system models: Novel methods and their impact on greenhouse gas emissions*, Journal of Industrial Ecology, 26 (2022), pp. 392–405.
- [16] T. WEI, X. CHU, D. YANG, AND H. MA, *Power balance control of res integrated power system by deep reinforcement learning with optimized utilization rate of renewable energy*, Energy Reports, 8 (2022), pp. 544–553.
- [17] H. ZHANG, *Key technologies and development challenges of high-proportion renewable energy power systems*, Highlights in Science, Engineering and Technology, 29 (2023), pp. 137–142.
- [18] Z. ZHANG, C. DOU, D. YUE, Y. ZHANG, B. ZHANG, AND Z. ZHANG, *Event-triggered hybrid voltage regulation with required bess sizing in high-pv-penetration networks*, IEEE Transactions on Smart Grid, 13 (2022), pp. 2614–2626.
- [19] K. ZHOU, Z. ZHANG, L. LIU, AND S. YANG, *Energy storage resources management: Planning, operation, and business model*, Frontiers of Engineering Management, 9 (2022), pp. 373–391.
- [20] K. ZUO AND L. WU, *Enhanced power and energy coordination for batteries under the real-time closed-loop, distributed microgrid control*, IEEE Transactions on Sustainable Energy, 13 (2022), pp. 2027–2040.

*Edited by:* B. Nagaraj M.E.

*Special issue on:* Deep Learning-Based Advanced Research Trends in Scalable Computing

*Received:* Aug 15, 2023

*Accepted:* Oct 16, 2023



## NUMERICAL SIMULATION AND OPTIMAL CONTROL OF COMPOSITE NONLINEAR MECHANICAL PARTS CASTING PROCESS

HUAN LI\* AND PENG WANG†

**Abstract.** In order to understand the numerical simulation and optimal control of the casting process of Machine element, the author proposed a study based on the numerical simulation and optimal control of the casting process of composite Machine element. Firstly, the author analyzed the structural characteristics of composite Machine element, introduced them in detail, and studied their casting technology in combination with the actual situation. Secondly, the casting process of composite Machine element was simulated by using numerical simulation method, and the temperature field and flow field in the casting filling and solidification process were analyzed. Finally, selecting square cylinders with different wall thicknesses as typical components, the scheme design of traditional single casting process and multi material composite casting process based on dieless casting composite forming technology were carried out. Finite element numerical simulation and experimental research were conducted on the two processes, respectively. The results indicate that: The casting obtained by the multi material composite casting process almost solidifies simultaneously around 200 seconds later, and the graphite morphology around the casting is Type A, with a length of about 100  $\mu\text{m}$ , with small differences and uniform distribution; The minimum difference in tensile strength around the casting is about 3.8%, and the maximum increase in tensile strength value is 21%. This research achievement can provide technical reference for high-performance and high-quality casting of complex iron castings. In order to improve the casting quality, the author optimized the casting process of composite Machine element by numerical simulation.

**Key words:** Composite materials, casting process, numerical simulation

**1. Introduction.** In the mechanical manufacturing industry, the application of composite material parts is becoming increasingly widespread [3, 10]. Due to the special properties of composite materials (such as dimensional stability, corrosion resistance, etc.) and their important role in mechanical manufacturing, composite parts are gradually replacing metal materials as an important component in the field of mechanical manufacturing. Composite material parts have many advantages compared to metal parts, such as light weight, high specific strength and modulus, good wear resistance, and strong designability. In addition, due to the differences in thermal physical and mechanical properties between composite materials and metal materials, the latent heat of phase change during the casting process is large, the heat transfer rate is slow, and the solidification time is long, these have had a significant impact on the casting process. Therefore, the casting process design of composite Machine element should fully consider the characteristics of large latent heat of phase change and slow heat transfer rate. Numerical Simulation and Optimization Control of Composite Machine element Casting Process During the casting process of composite Machine element, due to the different thermophysical and mechanical properties of composite and metal materials, the process of heat conduction and heat conduction during the solidification process is more complex. The use of numerical simulation technology can effectively predict and optimize casting processes, thereby shortening design cycles and reducing production costs. The mold filling and solidification process in the casting process has a great impact on the quality of the casting. For composite Machine element, the main component of the composite material is resin, and its latent heat of phase change is very large. If the composite parts are not pre treated (such as heat treatment, adding alloy elements, etc.) before pouring, it will lead to internal defects in the casting. In addition, due to the significantly lower thermal conductivity and thermal conductivity of composite material parts compared to metal parts, local overheating of the solidification site is prone to occur during the casting process. This Casting defect can be effectively prevented and controlled by reducing the filling time, shortening the solidification time and using materials with low thermal conductivity. The optimization design of casting process is to determine the optimal process plan based on the requirements of casting material performance, casting size,

---

\*Xinxiang Vocational and Technical College, Xinxiang, Henan, 453006, China (Corresponding author: [HuanLi32@163.com](mailto:HuanLi32@163.com))

†Xinxiang Vocational and Technical College, Xinxiang, Henan, 453006, China (Email: [PengWang2@126.com](mailto:PengWang2@126.com))

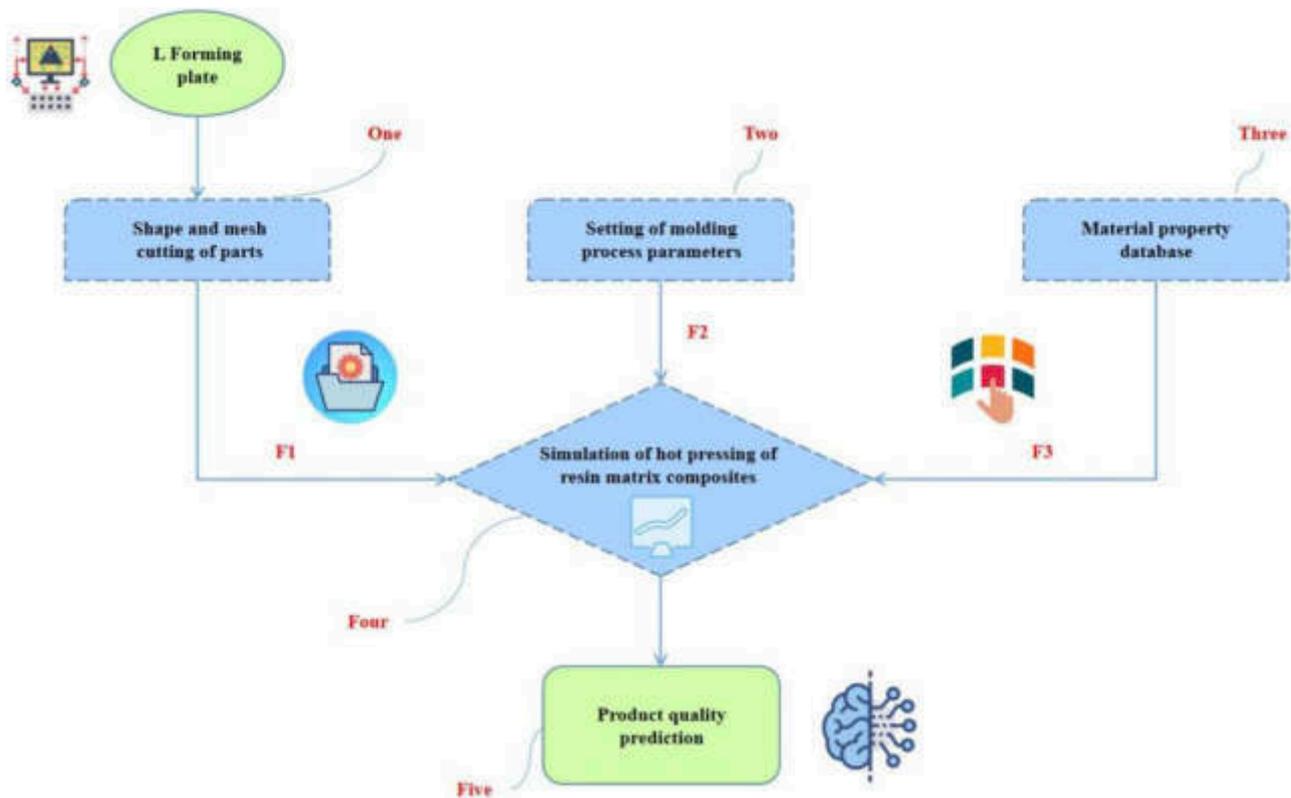


Fig. 1.1: Numerical simulation and optimization of composite machine element casting processflow

and casting structure; And the optimization design of casting process is an essential and important link in the casting production process, which is an important means to reduce production costs, improve product quality, improve labor productivity, and enhance the competitiveness of enterprises (Figure 1.1).

**2. Literature Review.** Machine element play a very important role in the production of modern enterprises, and composite Machine element are a typical composite structure, which have the advantages of light weight, high strength, corrosion resistance and good processing performance, and play a very important role in modern machining. Composite materials have the characteristic of anisotropy, and suitable casting processes need to be used during the processing to ensure product quality. At present, composite Machine element are mainly produced with resin matrix composites and Metal matrix composite. Resin based composite materials have the advantages of light weight, high strength, corrosion resistance, and good processing performance, making them a very ideal material. However, due to their anisotropic characteristics, traditional casting processes are difficult to ensure product quality in actual production processes. At present, the casting process numerical simulation of composite Machine element mainly uses computer numerical simulation technology to analyze the casting solidification process, and then optimize the casting process plan. By utilizing computer numerical simulation technology, it is of great significance to effectively predict and optimize casting process plans, thereby effectively improving product quality. Composite materials have the advantages of light weight, high strength, corrosion resistance, and good processing performance, which play a very important role in modern mechanical processing. However, composite materials have the characteristics of anisotropy, so suitable casting processes need to be adopted in production to ensure product quality. Composite Machine element are composed of metal materials and non-metallic materials, their interior is composed of multiple layers, and they also have certain anisotropy in structure, at the same time, because there are many pores in the interior, they are prone to appear shrinkage, porosity and porosity defects when casting them This is mainly due to the strong orientation

of fibers in composites. When Machine element are in a certain temperature environment, fibers will undergo anisotropic deformation along the fiber direction. This is mainly because there are a large number of pores in the composite material, and during the casting process, due to the presence of many internal pores, it is easy to have defects such as uneven solidification shrinkage and difficulty in demolding the core during pouring. Composite Machine element also have anisotropy and isotropy, which have a great impact on the casting process. Therefore, in actual production, it is necessary to control and optimize it by adopting reasonable process measures. The application of numerical simulation technology in the casting process of composite Machine element plays a very important role in improving product quality and reducing production costs. Due to the complex structure, small thickness and high density of composite Machine element, therefore, multiple factors need to be considered during the design process, which increases the difficulty of the design. During the casting process, due to the high temperature of the molten metal and the anisotropic nature of the material, defects such as shrinkage cavities and porosity are prone to occur. And numerical simulation technology can predict the defects that occur during the casting process before casting, analyze the defects to determine the defect location and cause of the casting, and then take corresponding measures for optimization treatment. Numerical simulation of composite Machine element can effectively reduce the occurrence of casting defects and improve product quality.

Li, J. et al. evaluated the molten pool size under different process parameters by numerical simulation and SEM analysis, and the results showed a good agreement between the two. In addition, the effect of HA on the mechanical properties of sls was also studied [8]. Mulikapuri, S. et al., through different controlled experiments, can infer that bismuth-aquatic complexes are functional sites in response to OER. The overpotential of pH 13.0, pH 7.0 and pH 4.0 solutions was 375.0 mV (relative to reversible hydrogen electrode (RHE)), 585.21 mV (relative to reversible hydrogen electrode) and 830.10 mV (relative to reversible hydrogen electrode), respectively. The catalytic conversion frequencies of bismuth-aquatic complexes were 1.24, 5.90 and 0.93-s-1, respectively [9].

### 3. Research Methods.

**3.1. Mold design method of typical multi material composite machine element.** Square cylinders with different wall thicknesses (8 mm, 10 mm, 15 mm and 18 mm) were selected as typical parts, which have uneven wall thicknesses and geometric thermal joints. Considering that the cast iron parts should approximately satisfy the principle of simultaneous solidification and the mechanical properties are expected to be uniform, the square barrel is split into six sand units. The sand setting of each sand unit was calculated, and four sand units with different wall thickness edges were obtained, and different types of sand were applied, and resin silica sand was used for the sand core unit and the base sand unit. As shown in equation (3.1)

$$t_f = (0.94 \sim 0.97) \frac{\pi \rho_s L^2 V^2}{4 \lambda_m c_m \rho_m A^2 (T_m - T_0)^2} = (0.94 \sim 0.97) \quad (3.1)$$

In the formula:  $T_m$  is the melting point of the metal casting,  $T_0$  is room temperature, °C;  $P_s$  is the density of the metal casting, kg/m<sup>3</sup>;  $L$  is the latent heat of solidification, W/m<sup>3</sup>;  $B_m$  is the heat storage coefficient of the mold.

### 3.2. Finite element numerical simulation of typical casting process.

**3.2.1. Establishment of finite element model for typical components and pre processing loading.** Using the casting/mold integration modeling method, first build the Geometric modeling of the cubic cylinder and its mold [6, 16]. And then divide the surface grid and volume grid of the mold and casting assembly through the MeshCAST function module in ProCAST software. This grid division mainly adopts tetrahedral grids, and local grid refinement has been applied to important parts.

The pre processing module Precast of ProCAST software sets the thermal properties parameters, gravity parameters, boundary conditions, interface heat transfer coefficients, and casting materials for the finite element models of each specimen, as shown in Table 3.1.

**3.3. Numerical simulation of temperature field during solidification process of typical components.** In this study, the temperature field of the traditional Machine element casting process and the composite Machine element casting process of the square cylinder were numerically simulated, respectively, to

Table 3.1: Pre processing conditions settings

Parameter Name	Design parameter settings
Casting material	Casting: Gray iron 260
Interface heat transfer coefficient	Reference self test results
boundary condition	Pouring temperature: 1460 °C
Gravity parameters	Direction: Gravity direction (vertically downwards)

obtain the solidification time (that is, all the liquidus time) [15, 12]. According to the numerical simulation results, there is a significant difference in the solidification time of the four sides of the square cylinder in traditional casting processes. The solidification time of the thinner edges reaching the liquidus is about 300 seconds, while the solidification time of the thicker edges is about 650 seconds, with a difference of about 50% [5, 14, 11]. The composite casting process uses chromite ore sand with stronger heat storage capacity for the thicker side, and uses the quartz sand with stronger heat storage capacity for the thickest side, so the solidification time of all sides of the casting is about 300 s, almost achieving simultaneous solidification [18, 4].

**3.4. Numerical simulation of metallographic structure and mechanical properties of typical parts.** This study conducted numerical simulations on the metallographic structure and mechanical properties of cast iron castings obtained from traditional and composite material casting processes of typical castings [20]. The thinnest edge of graphite volume fraction is about 0.01, the thickest edge is about 0.0137, and the tensile strength value is about 420 MPa for the thinnest edge and 360 MPa for the thickest edge. The graphite volume fraction around the square cylinder obtained by the composite material casting process is about 0.133, about 420 MPa. From the numerical simulation results, it can be seen that there is a significant difference in graphite volume fraction and tensile strength values around typical castings in traditional casting processes, and the tensile strength value is almost the same, indicating that the overall mechanical properties of the typical component are uniform and good.

#### 4. Experimental Analysis.

**4.1. Typical casting process test steps and conditions.** The ratio of various types of sand is: alkaline phenolic resin accounts for 3% of the weight of the original sand, and curing agent accounts for 30% of the weight of the resin. Cut and process each sand mold unit on a digital dieless casting precision forming machine, and assemble each sand mold unit into a mold to be poured [1]. When the temperature of the molten iron reaches 1460 °C, pouring is carried out to obtain a square cylinder casting.

**4.2. Temperature field test for solidification of square cylinders.** The real-time temperature value during the solidification process of the square cylinder measured by a high-temperature thermocouple embedded in the mold is shown in Figure 4.1(a) (b).

According to the results, the temperature field test results of the solidification process of square simple parts are consistent with the finite element numerical simulation results: In traditional casting techniques, the solidification time of square cylindrical parts varies greatly from the beginning of solidification to the liquidus. The solidification temperature gradient of the thinner edge is larger, while the solidification temperature gradient of the thicker edge is smaller. At 100 seconds, the solidification temperature difference at different wall thicknesses is 80 °C. In the composite casting process, there is a slight difference in temperature gradient at the beginning of solidification on the four sides of the square cylinder. At 100 seconds, the solidification temperature difference at different wall thicknesses is 45 °C, and after 210 seconds, the temperature gradient is almost the same, achieving simultaneous solidification.

**4.3. Metallographic structure testing of square cylinder parts.** Cut and grind typical pieces obtained from traditional casting processes and composite material casting processes, and observe the graphite distribution morphology and length under a high-power optical microscope.

During the solidification process of castings, the cooling rate depends on the thickness of the casting wall, and there is a significant difference in the cooling rate between the thick walled and thin-walled parts of the

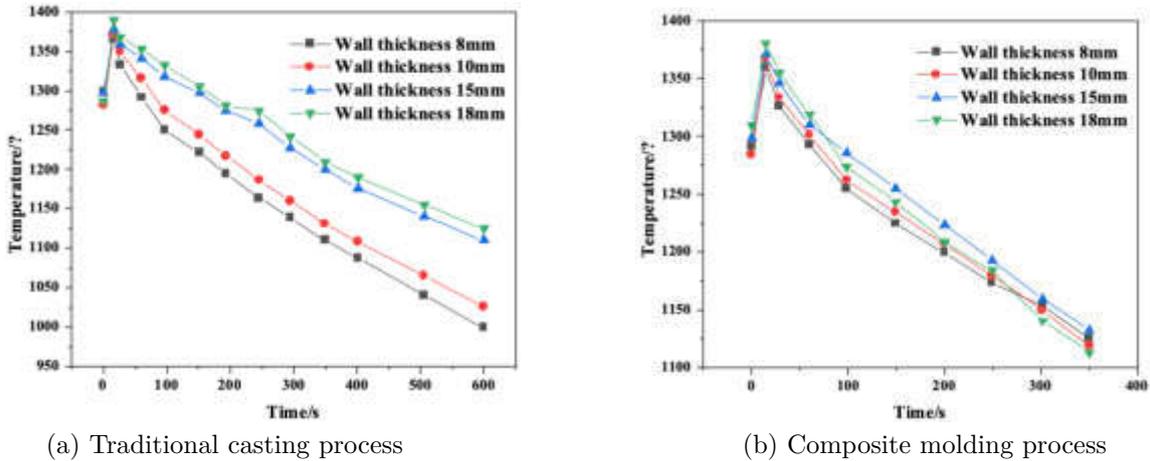


Fig. 4.1: Measured temperature during the solidification process of a square cylinder

Table 4.1: Tensile strength of square cylinders

Wall thickness /mm	Tensile strength of typical components in traditional casting processes/MPa				Tensile strength of typical components in load material casting process /MPa			
	1	2	3	average	1	2	3	average
9	303	310	321	311.33	310	304	315	309.66
12	285	294	300	293	306	308	302	305.33
16	242	260	250	250.66	297	300	294	330.33
20	240	245	244	243	289	285	290	288

casting. The graphite morphology of parts with significant differences in wall thickness of the same casting may also vary. For example, traditional casting processes have significant differences in graphite morphology and length around castings with different wall thicknesses. The graphite morphology includes A type and A+E type, with a length of 100-300  $\mu\text{m}$ . The composite material casting process uses different types of molding sand with different heat storage capabilities for different wall thicknesses, resulting in a graphite shape around the casting that is similar to Type A, with a length of about 100  $\mu\text{m}$ , with small differences and uniform distribution.

**4.4. Mechanical property testing of square cylinders.** The tensile strength of typical specimens cut from traditional casting and composite material casting processes is tested on a universal material testing machine, as shown in Table 4.1 and Figure 4.2.

According to the test results, the tensile strength values of the four sides of the square cylinder obtained by the traditional casting method are different, the maximum value is 311.33 MPa, and the minimum value is 240 MPa. A difference of about 20. Approximately less, the maximum value is 330.33 MPa, and the minimum is 285 MPa, with a difference of about 4. This indicates that the overall mechanical properties of the typical component are approximately uniform and consistent, and the maximum increase in tensile strength value of the casting is 21%. The reason for this is that the graphite size around the square cylinder obtained from the composite material casting process is small, similar in length, and evenly distributed, resulting in an increase in the tensile strength of the thick wall and a small difference in the overall strength of the casting. In addition, there are some differences between the tensile strength test results of typical components and the strength values of finite element numerical simulation results, but it has no substantial impact on the test results [17, 19, 2, 13, 7]. This is because the performance parameters of the HT250 material used for pouring after inoculation treatment

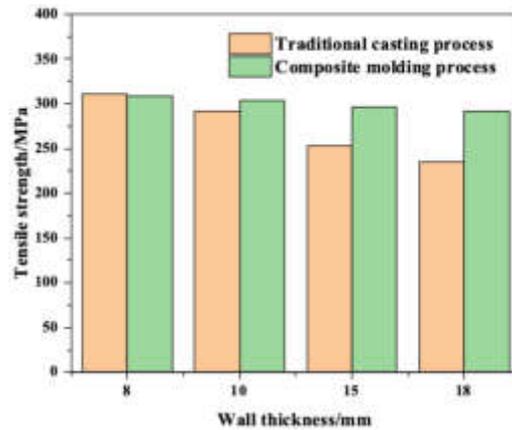


Fig. 4.2: Comparison of tensile strength of square cylindrical components with different wall thicknesses

differ from the performance parameter values of HT250 in the ProCAST software material library.

**5. Conclusion.** (1) In traditional casting processes, the solidification time of all four sides of a square cylinder reaching the liquidus varies greatly. The thinner edge solidifies first, while the thicker edge solidifies last. At 100 seconds, the temperature difference between different wall thickness edges is about 96 °C. In the composite material casting process, there is a slight difference in temperature at the beginning of solidification on the four edges of the square cylinder. At 100 seconds, the temperature difference between different wall thickness edges is about 35 °C, and after 200 seconds, almost simultaneous solidification is achieved.

(2) The graphite morphology and length around square tube castings with different wall thicknesses obtained by traditional casting techniques vary greatly. The graphite morphology includes A type and A+E type, with a length of 100-300  $\mu\text{m}$ . The graphite morphology around the castings obtained by the composite casting process is all A-type, with a length of about 100  $\mu\text{m}$ , with small differences and uniform distribution.

(3) The tensile strength values of the four sides of the square cylinder cast by traditional Machine element differ greatly by about 24%. The difference in tensile strength values around the square cylinder obtained from the composite material casting process is small, with a difference of about 3.8%. The overall mechanical properties of this typical component are approximately uniform, and the maximum increase in tensile strength value of the casting is 21%.

## REFERENCES

- [1] A. F. ABATE, L. CIMMINO, I. CUOMO, M. DI NARDO, AND T. MURINO, *On the impact of multimodal and multisensor biometrics in smart factories*, IEEE Transactions on Industrial Informatics, 18 (2022), pp. 9092–9100.
- [2] F. CALIGNANO, F. GIUFFRIDA, AND M. GALATI, *Effect of the build orientation on the mechanical performance of polymeric parts produced by multi jet fusion and selective laser sintering*, Journal of Manufacturing Processes, 65 (2021), pp. 271–282.
- [3] V. CHAUDHARY AND P. P. GOHIL, *Stress analysis in cotton polyester composite material*, International Journal Metallurgical & Materials Science and Engineering, 3 (2013), pp. 15–22.
- [4] E. FRACCHIA, F. S. GOBBER, C. MUS, Y. KOBAYASHI, AND M. ROSSO, *About residual stress state of castings: The case of hpdc parts and possible advantages through semi-solid processes*, Solid State Phenomena, 327 (2022), pp. 272–278.
- [5] A. GANAPATHI ET AL., *Se-doped sio2 nanocomposite material synthesis, characterization and multi applications*, Materials Science for Energy Technologies, 5 (2022), pp. 161–170.
- [6] D. A. KOLOSOV AND O. E. GLUKHOVA, *A new composite material on the base of carbon nanotubes and boron clusters b12 as the base for high-performance supercapacitor electrodes*, C-Journal of Carbon Research, 7 (2021), p. 26.
- [7] L. LE, M. A. RABSATT, H. EISAZADEH, AND M. TORABIZADEH, *Reducing print time while minimizing loss in mechanical properties in consumer fdm parts*, International Journal of Lightweight Materials and Manufacture, 5 (2022), pp. 197–212.
- [8] J. LI, Y. YANG, Z. ZHAO, AND R. YAN, *Numerical modeling and experimental verification of the composite material of pa12/ha with parameterized random distribution in selective laser sintering*, Rapid Prototyping Journal, 28 (2022), pp. 873–883.

- [9] S. MULKAPURI, A. RAVI, AND S. K. DAS, *Fabricating a functionalized polyoxometalate with zif-8: a composite material for water oxidation in a wide pH range*, Chemistry of Materials, 34 (2022), pp. 3624–3636.
- [10] D. I. PERMATASARI, B. RUSDIARSO, AND N. NURYONO, *Green synthesis and characterization of natural magnetic particles/chitosan composite material impregnated with copper nanoparticles*, Solid State Phenomena, 339 (2022), pp. 19–27.
- [11] M. F. QI, Y. L. KANG, Y. H. ZHENG, J. C. WANG, G. N. LI, AND Y. ZHANG, *Industrialization of aluminum alloy uniform solidification controlled rheological die casting*, Solid State Phenomena, 327 (2022), pp. 163–171.
- [12] P. O. RUSINOV AND Z. M. BLEĐNOVA, *Study of the structure and properties of a high-entropy ceramic composite material*, Surface Innovations, 10 (2021), pp. 217–226.
- [13] E. SALLICA-LEVA, F. H. DA COSTA, C. T. DOS SANTOS, A. L. JARDINI, J. V. L. DA SILVA, AND J. B. FOGAGNOLO, *Microstructure and mechanical properties of hierarchical porous parts of ti-6al-4v alloy obtained by powder bed fusion techniques*, Rapid Prototyping Journal, 28 (2022), pp. 732–746.
- [14] S. SINGH, R. KUMAR, R. KUMAR, J. S. CHOCHAN, N. RANJAN, AND R. KUMAR, *Aluminum metal composites primed by fused deposition modeling-assisted investment casting: Hardness, surface, wear, and dimensional properties*, Proceedings of the Institution of Mechanical Engineers, Part L: Journal of Materials: Design and Applications, 236 (2022), pp. 674–691.
- [15] X. WANG, Y. LIANG, S. WEI, Y. WANG, Y. YUAN, L. LI, B. WANG, AND H. XIE, *Preparation and absorbing property analysis of zno-doped magnetic graphene composite material*, Journal of Magnetism and Magnetic Materials, 556 (2022), p. 169450.
- [16] M. YE, Y. TENG, X. ZHAO, S. WANG, J. TANG, AND R. AHUJA, *Preparation and characterization of sic-zac composite material for immobilizing the simulated radioactive graphite*, Progress in Nuclear Energy, 143 (2022), p. 104029.
- [17] S. ZHANG, Y. XU, W. ZHANG, AND X. HUI, *Micro-mechanical modeling study of the influence of cure process on the interfacial cracking of z-pinned laminates*, Composite Structures, 280 (2022), p. 114889.
- [18] Y. ZHAO, J. BI, Z. WANG, L. HUO, J. GUAN, Y. ZHAO, AND Y. SUN, *Numerical simulation of the casting process of steel fiber reinforced self-compacting concrete: Influence of material and casting parameters on fiber orientation and distribution*, Construction and Building Materials, 312 (2021), p. 125337.
- [19] L. ZHU, Z. YANG, B. XIN, S. WANG, G. MENG, J. NING, AND P. XUE, *Microstructure and mechanical properties of parts formed by ultrasonic vibration-assisted laser cladding of inconel 718*, Surface and Coatings Technology, 410 (2021), p. 126964.
- [20] D.-D. ZHUANG, B. DU, S.-H. ZHANG, W.-W. TAO, Q. WANG, AND H.-B. SHEN, *Effect and action mechanism of ultrasonic assistance on microstructure and mechanical performance of laser cladding 316l stainless steel coating*, Surface and Coatings Technology, 433 (2022), p. 128122.

*Edited by:* B. Nagaraj M.E.

*Special issue on:* Deep Learning-Based Advanced Research Trends in Scalable Computing

*Received:* Aug 25, 2023

*Accepted:* Oct 16, 2023



## CONSERVATION DESIGN OF INDUSTRIAL HERITAGE BASED ON NONLINEAR GA OPTIMIZATION ALGORITHM AND THREE-DIMENSIONAL RECONSTRUCTION

YUNAN ZHAO\* AND PENG BAI†

**Abstract.** In order to understand the industrial heritage protection design of Iterative reconstruction, the author proposes a research on industrial heritage protection design based on GA optimization algorithm and Iterative reconstruction. Firstly, the author establishes the 3D model of industrial heritage through Iterative reconstruction, and optimizes the model parameters through GA algorithm to achieve the purpose of protecting and utilizing industrial heritage. Secondly, the author proposes a method of Iterative reconstruction of industrial heritage based on GA algorithm, uses this method to conduct Iterative reconstruction of industrial heritage, and imports the reconstructed model into the 3D model management system for management. This method solves the problem of high reconstruction cost caused by low model quality in traditional Iterative reconstruction, and makes industrial heritage protection design more practical. Finally, an experimental analysis was conducted using a factory building in a certain city as an example. The results showed that the model optimized using the GA algorithm had significantly better performance than traditional reconstruction methods, and could more accurately reflect the spatial form and structural characteristics of industrial heritage, this provided new ideas and methods for the subsequent protection and utilization of industrial heritage. The GA algorithm optimized 3D model established by the author can effectively evaluate industrial heritage in historical urban areas, not only revealing the value of industrial heritage better, but also providing a certain reference for similar work in the future.

**Key words:** GA optimization algorithm, 3D reconstruction, industrial heritage protection design

**1. Introduction.** With the acceleration of economic development and urbanization, the rapid development of cities, and the shortage of urban land, the number of industrial heritage in cities has been increasing year by year, and its quantity is also constantly increasing. At present, China has carried out preliminary protection and utilization of industrial heritage. However, there are certain misunderstandings in the protection and utilization of industrial heritage, mainly manifested in: firstly, emphasizing the value of industrial heritage itself unilaterally while neglecting its social, economic, and cultural value; Secondly, simply redeveloping and utilizing industrial heritage, neglecting the relationship between its historical value and artistic value; Thirdly, it is believed that protection is the demolition and reconstruction of historical buildings, without realizing that protecting industrial heritage requires a comprehensive consideration from a macro perspective.

When protecting and utilizing industrial heritage, attention should be paid to a comprehensive evaluation of its historical, artistic, and socio-economic value. However, due to the long history of industrial heritage, the rich historical information and cultural value it contains often cannot be expressed in language or evaluated through traditional methods. Therefore, the author mainly studies how to use GA algorithm for three-dimensional modeling of industrial heritage in historical urban areas. GA algorithm is a global optimization algorithm based on Evolutionary computation, it solves optimization problems under multi-objective and multi constraint conditions by simulating biological genetic processes. In the GA algorithm, the most important aspects are the design of population size and selection operators. The larger the group size, the more optimal solutions will be obtained; The selection operator determines the direction and scope of group search. Therefore, selecting the appropriate population size and selecting the appropriate selection operator are the two most important issues in optimization algorithms. At present, GA algorithm is commonly used to solve nonlinear problems in optimization problems. The GA algorithm is divided into two stages: encoding stage and fitness function calculation stage. The fitness function calculation stage is divided into two types: Global optimization and local optimization. In addition, the GA algorithm can also be used to solve nonlinear problems (such as regression and clustering analysis) [1, 17] (Figure 1.1).

---

\*Shenyang Normal University, College of Fine Arts Design, Shenyang, Liaoning, China (YunanZhao6@163.com)

†Shenyang Normal Univ., College of Fine Arts&Design, Shenyang, Liaoning, China (Corresponding author, PengBai26@126.com)

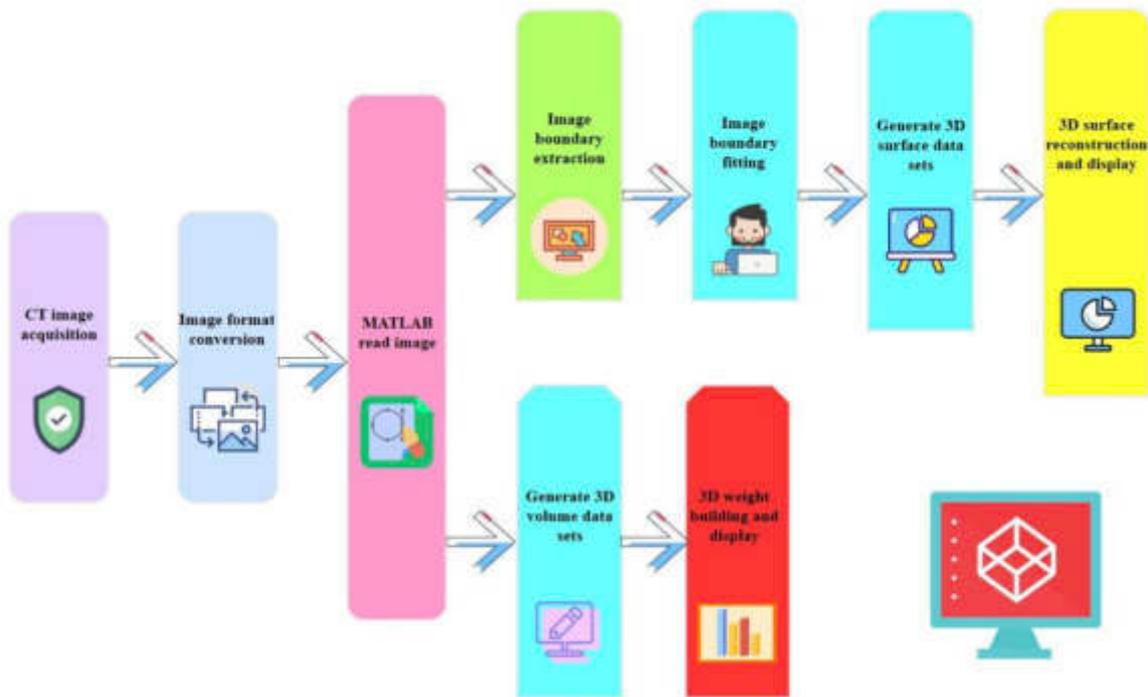


Fig. 1.1: Iterative reconstruction of industrial heritage protection design

**2. Literature Review.** Industrial heritage is a special product generated in the process of economic globalization and urbanization, which has a significant impact on urban development. In recent years, with the continuous renewal and renovation of cities, many industrial heritage sites have been abandoned. At the same time, people are paying increasing attention to the protection and reuse of industrial heritage, and research on its protection and reuse is also increasing. However, there are still some problems in the current industrial heritage protection design, such as the low quality of the model in the Iterative reconstruction process, the lack of a unified management system, and so on. These issues have had a negative impact on the protection and reuse of industrial heritage. In order to solve these problems, the author proposes a method of Iterative reconstruction of industrial heritage based on GA algorithm. First, we use computer graphics to model industrial heritage; Secondly, import the model into the 3D model management system for management; Finally, the GA algorithm is used to optimize the model parameters to achieve better modeling results. Industrial heritage is a product of industrial civilization, carrying the memory of a city and witnessing its history. However, due to the fact that most of the industrial heritage is old factory buildings, which themselves have high value, but due to their own characteristics, they are not suitable for development as tourist attractions. In recent years, many scholars in China have also conducted relevant research on industrial heritage. Lyu, S. Yangshan "ceramic industrial site as an example, on the basis of field research and photography of the site, we used Photoshop and 3 dsMax software to conduct 3D modeling of the site, and completed the Iterative reconstruction of the site [8]. Li, C., and others used laser scanners to measure and model the No.1 and No.2 blast furnaces of China First Steel Company, and visualized the models using VS2010 programming language [5]. Teng, X. and others proposed an industrial heritage protection plan by analyzing the historical value of China's First Steel Company's No. 2 blast furnace [12]. In the traditional Iterative reconstruction method, taking 3DMAX as an example, its parameter setting mainly includes point cloud data acquisition, point cloud data processing, curve generation, surface generation, etc. However, this method takes a long time and is inefficient. For this reason, the GA based optimization algorithm is applied to Iterative Reconstruction of Industrial Heritage. After modeling, we need to manage and query the model, which is convenient for the user to manage and utilize the model. To

achieve the Iterative Reconstruction of Industrial Heritage, the first step is to construct a model. The usual way is to construct the model based on the structure features. When building a model, the resulting 3D model will be affected differently by different parameters. Therefore, it is necessary to optimize the model parameters for Iterative reconstruction of industrial heritage [13].

### 3. Research Methods.

**3.1. A nonlinear optimization design method based on GA.** CA is aimed at unconstrained problems, and for dynamic optimization design problems that are generally constrained optimization, conversion is required. Therefore, the dynamic optimization design problem is modified as follows:

Construct a penalty function as shown in equation (3.1):

$$F(b^j) = \Psi_0 + S \sum_{i=1}^m \max\{0, g_i\} \quad (j = 1, 2, \dots, N) \quad (3.1)$$

$g_i$  is the  $i$ -th constraint function, as shown in equation (3.2):

$$g_i(t, b, z) \leq 0 \quad (0 \leq t \leq T, i = 1, 2, \dots, m) \quad (3.2)$$

$S$  is a large positive number.

Constructing Individual Moderate Function Equation (3.3):

$$G(b^j) = \frac{1}{|F(b^j) - c|} \quad (3.3)$$

where  $c$  is a positive number that is small enough.

The characteristic of dynamic optimization design is that the objective function and constraint conditions are functions of system parameters and system dynamics equation solutions. During the optimization process, numerical solutions of the system dynamics equation must be performed for each optimization. However, in practical applications, the general GA algorithm exposes the shortcomings of slow progress and multiple calculations of the evaluation function. The main reason for this is that there is too much randomness in the seed selection process, which covers the entire solution space. However, the overall quality is very poor, and many are even infeasible. It takes several generations of genetic iteration to improve the overall quality. Therefore, improvements must be made in order to be used for dynamic optimization design. In practical dynamic optimization design problems, they are often based on their original solutions, based on this characteristic, when forming the original group, the original plan is imposed as a part of the original group, and after a few iterations, the overall quality is rapidly improved [20, 10]. In addition, the composite method can also be used to optimize seed selection, which is described as follows:

Evaluate a group of schemes including  $k$  individuals, select the worst-case individuals, and modify them as follows: Take each individual in the group scheme except for the worst-case individual as a vertex of the polygon, and the centroid position of the polygon is expressed in equation (3.4)

$$x^r = x^c + \alpha(x^c - x^h) \quad (3.4)$$

By taking the above measures, the optimization process was accelerated and the optimization time was shortened, the modified algorithm is as follows:

- (1) At the beginning ( $k=0$ ), evenly divide the solution space of each design variable, randomly select a set of original solutions from these numerous small domains, and impose the original solution as a part of the original population based on the actual optimization problem,
- (2) Solve the dynamic equation and obtain the penalty function  $F(B_j)$  and fitness function  $G(b_j)$ .
- (3) Selection: Based on the probability of inclusion in the original solution.
- (4) Randomly select two independent individuals (including bm and bI) from the excellent varieties obtained in the previous step to hybridize and generate new individuals.
- (5) Perform genetic mutations on individuals with  $P_j \leq P_m$  ( $P_m$  is the probability of variation) and individuals with similar quality in the original solution to generate new individuals.

Table 3.1: Comparison table

	$K_1/\text{KN} \cdot \text{m}^{-1}$	$K_2/\text{KN} \cdot \text{m}^{-1}$	$K_3/\text{KN} \cdot \text{m}^{-1}$	$K_4/\text{KN} \cdot \text{m}^{-1}$
GA algorithm	8.75	35.13	42.18	1.95
traditional algorithm	8.72	36.00	43.001.95	2.36

(6) Form a new generation population of the same size as the original by combining the best and new individuals, and repeat steps (2) to (5) until the average moderate stabilizes.

In order to verify the effectiveness of the above methods, four optimization design problems for the five degree of freedom vehicle suspension system are selected as examples in this paper, as shown in Table 3.1.

The numerical example shows that the GA based dynamic optimization design method proposed by the author is feasible, and careful analysis shows that it has the following characteristics compared to other dynamic optimization design methods:

- (1) Sensitivity analysis can be omitted during the optimization process;
- (2) Minimal requirements for the objective function and constraints, and no feasibility requirements for the initial solution;
- (3) The optimization search covers the entire solution space and performs heuristic searches based on genetic mechanisms, resulting in a high probability of obtaining the global optimal solution;
- (4) It is equally effective for both nonlinear and linear problems.

**4. Experimental Application of Iterative Reconstruction Industrial Heritage Protection and Renewal Design in a Factory.** The author conducted a case study on the analysis of a factory area and the reconstruction of building environmental information models, vector data reconstruction and building performance analysis (building elevation, building shadow, solar radiation) were conducted on the factory, and the industrial heritage of the factory was updated and designed, including proposing update goals, protecting update positioning, and updating design strategies [19, 6].

#### 4.1. Objectives and principles of protection and renewal.

**4.1.1. Protection and update objectives.** (1) Inject new vitality and strengthen the protection of the old factory. In the development of the city, the old industry is slowly disconnected from the development of the city, and the society and individuals have a weak awareness of the protection of industrial heritage. Industrial heritage is an important part of the social Collective memory, an inseparable culture of a city, and has a special status. The 156 projects in Xiangfang District are more important industrial plans after the founding of New China, which have important historical research value and technical value. These factories' construction techniques represent the advanced construction techniques and structure standards. In the meantime, these enterprises have taken in a lot of city residents, and have close relationship with city life and work. In recent years, because of the relocation of some factories, they have been abandoned, and many factories are facing the danger of being demolished. This paper discusses how to inject new elements into the old factory, change its functions and revitalize it, and make use of "active protection" instead of "static protection", so as to make full use of industry heritage and enhance the protection of industrial heritage.

(2) Improve the level of information protection of industrial heritage This paper explores the use of building information Iterative reconstruction technology for digital information collection of industrial heritage. With the advantages of digital information technology that is easy to preserve, process, observe, and analyze, 3D measurement and data processing of industrial heritage and the preservation of 3D data of building information in the original factory area are carried out. At present, the protection of industrial heritage faces great difficulties in the evaluation and data acquisition of architectural heritage. Traditional data acquisition and recognition models have significant limitations. The acquisition of traditional manual two-dimensional data makes it difficult for the data to intuitively display the information of the factory area, at the same time, there are certain errors that consume a lot of manpower and resources. Utilizing digital information technology in architecture to conduct data collection and auxiliary design planning for the factory will improve work efficiency and provide a demonstration role for data collection and planning of other industrial heritage sites [9, 14].

**4.1.2. Protection and renewal principles.** (1) The principle of historical inheritance is that the first thing to be done in the protection of industrial heritage is to protect historical heritage information. As a carrier of historical information, the protection and updating of industrial heritage should first respect the information of the original buildings, maintain representative elements, buildings, and spatial layout, reflecting the unique industrial culture of the city, when injecting new elements into the existing factory area, attention should be paid to the protection of existing information. For industrial buildings, attention should be paid to the protection of information related to the style and characteristics of the factory area, for example, attention should be paid to the protection of the facade decoration of the main buildings, the overall structure of the plant, the production facilities of representative production processes, and the overall spatial structure of the plant area. These elements constitute the urban Collective memory of industrial buildings, which play a vital role in the protection of industrial heritage, continue the history and culture of the plant, and rely on the existing buildings to establish a complete system for the protection of industrial heritage [15].

(2) The design of the urban sharing principle should combine the characteristics of Xiangfang District, aiming at connecting the original factory with the city, taking the factory as the test point, gradually combining with other old industrial sites in Xiangfang District, and ultimately serving the residents of the entire urban area, at the same time, it forms a brand effect, promotes tourism, drives the development of the entire region, generates positive feedback, and the positive effect is feedback on the protection of industrial heritage. Form a “people-oriented” space to enhance the urban service level of the entire region.

**4.1.3. Evaluation of the building value of the factory.** The process of industrial heritage protection and subsequent development and utilization can be summarized into three processes: investigation, evaluation, and decision-making. Among them, the evaluation of the value of industrial heritage is an important link. The evaluation of industrial heritage value is the evaluation of its value, as well as the evaluation of its preservation status and future development potential. The evaluation of its value through quantitative and qualitative methods can effectively propose decision-making effects for subsequent protection and updates.

(1) The evaluation factors and levels of industrial heritage are due to the weak awareness of protecting industrial heritage among most people, and the lack of strong policy and regulatory support for industrial heritage. Some valuable industrial heritage has been demolished, causing irreparable losses. Based on the Wuxi Recommendation and the Beijing Initiative, the country passed the Interim Measures for the Management of National Industrial Heritage in 2018, emphasizing the importance of industrial heritage protection and hoping to incorporate industrial protection into the government’s urban planning and provide special funding support. And strengthen the publicity and reporting of national industrial heritage, and use Internet Big data, cloud computing and other means to create industrial art works; It is advocated to create Industrial tourism routes with popular science education, leisure and entertainment functions by focusing on “activation and protection” of industrial heritage and taking industrial heritage as the theme.

According to relevant domestic regulations, value evaluation can be divided into five categories, including historical value, technological value, social value, artistic value, and reuse value, these five categories are divided into 20 sub categories, and each sub category is divided into three grades, namely 0, 1 and 2 points. A score of 0 indicates that the item does not have the value, 1 indicates that it has general value, and 2 indicates that it has outstanding value [16, 18]. Divide the results into three levels: 1) Objects with special value can be protected with emphasis, maintaining their original characteristics without changing their characteristics and forms. The new use must be in line with the spirit of the original venue, and can be used as exhibition halls, exhibition halls, and other property buildings, and should not be developed as commercial properties. 2) By making appropriate changes to objects with general value, while protecting the appearance, structure, and landscape characteristics of the site, the function can adapt to the later functional needs. It can be transformed into a building type mainly focused on office, commercial, and tourism. 3) Buildings that do not require protection can be demolished or renovated according to the needs of later renovation. The evaluation criteria are shown in Table 4.1.

(2) Evaluation of the Building Value of the Factory - After investigation, we have classified the buildings within the factory area into production buildings, research office buildings, and ancillary buildings, an analysis and comparison were conducted on the historical value, social value, technological value, artistic value, and reusable value of buildings, and suitable buildings for preservation and buildings for reuse after renovation

Table 4.1: Industrial heritage evaluation factors

Value	Evaluation factor
Technological value	Style (architecture, environment) era
social value	pioneering
artistic value	Regional Characteristics and Aesthetic Characteristics of the Era
Reuse value	Physical availability, ecological value, location value

were classified, providing a basis for the planning, updating, and design of the later factory area. The site is roughly rectangular in shape. Within the site, the buildings are arranged in a regular manner, with production buildings occupying the main area of the factory. The main entrance is located on the side of Heping Road, and buildings 1 (the main building of the factory) and 2 (the cafeteria) are adjacent to the main entrance. There are some open spaces in the site planted with trees, including a small number of temporary factories [3].

**4.2. Update and optimization of the industrial and overall layout of the factory park.**

**4.2.1. Subdivision of park industries.** The construction of creative parks is not a traditional commercial real estate development project, but rather an industrial value operation based on industrial value space. The location and development positioning of creative parks are key factors for their successful operation, when positioning the industrial functions of the park, it is necessary to fully consider the local resource advantages and industrial foundation, avoid blindly following the trend and choosing industries without characteristics. When designing protective updates for industrial heritage, it is necessary to consider the resources, market, location, transportation, talent, and other aspects of the region. The era of large-scale industrial development has passed, and more detailed industrial segmentation has become the main development mode of contemporary times.

(1) The city has a long history of entertainment consumption. In the late 19th and early 20th centuries, with the construction of the Middle East railway, industry, commerce, and population gathered here. A city mainly focused on commercial port trade gradually formed, and the development of the city made commerce in the city prosperous. It is a popular tourist city and an international ice and snow city in the Northeast region, known as the “Little Paris of the East”.

(2) Exhibition function

Utilizing existing factories and open spaces can form exhibition spaces, which can provide space for enterprises or individuals to hold exhibitions and promote the cultural attributes of the park. The factory buildings in the park have the characteristics of large space and stable structure, similar to all building renovations. In the process of reusing the factory buildings, it is necessary to grasp their main characteristics and original elements that can represent their own value, during the renovation process, it is necessary to fully utilize the original structure and facade image of industrial buildings, preserve the original facade features, repair or replace some dilapidated components, and strive to reflect the value of their industrial heritage in the renovation results, so as to preserve the beauty of the original industrial buildings [7].

(3) Research Office

As the driving force of regional innovation, scientific research office can provide long-term driving force for the park and even Xiangfang District. At present, Xiangfang District, as a development area based on the old industrial zone, has gradually lagged behind other regions in terms of industrial production mode. The lack of investment in innovative research and development has slowed the economic development of the region. The introduction of scientific research and office work can promote the economic transformation of Xiangfang District, and is conducive to the diversified development of the park.

As shown in Table 4.2, for the renovation of the factory, a music area, exhibition area, creative office area, entertainment experience area, and commercial catering area will be set up. The functions of each partition are shown in the table below.

**4.3. Digital 3D management of the park.** The intelligent and information-based construction of creative parks combines the construction of the park with building 3D information models and 5G networks,

Table 4.2: Planning of park functions and content

Serial Number	Major function	Content planning
1	Musics	Music theater, music themed skits
2	Exhibition area	Cultural display, historical and cultural architecture display
3	Creative Office	Home of Creativity, Striving for the World, Cultural Education
4	entertainment experience	Catering, Fashion Release, Entertainment City

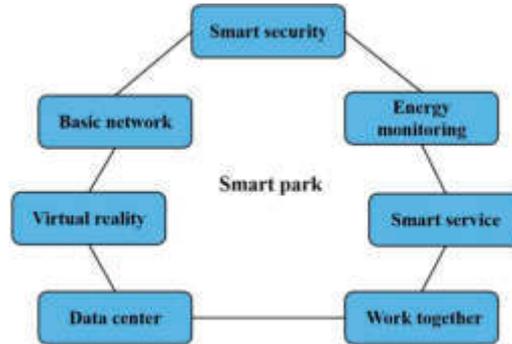


Fig. 4.1: General layout after renovations

comprehensively improving the level of informatization, intelligence, and integration of the park, and enhancing its visibility; In terms of personnel access management, vehicle management, and material distribution, the combination of virtual reality platforms and networks enables real-time monitoring and management, achieving a safe and convenient office environment, and achieving a scientific management platform that is intelligent, personalized, and efficient.

The construction of smart parks is based on building environmental information models, utilizing internet connectivity to transform the operation and management of the park from a decentralized model to an intensive development model. Based on a large cloud computing platform, data is shared, and the operation of the park is connected with the government and enterprises, achieving a multi-party collaborative construction pattern. See Figure 4.1.

By combining photogrammetric models with GIS, a three-dimensional circular system is established that integrates data collection, management, maintenance, update, and application, achieving the transformation of park management from two-dimensional to three-dimensional management, reducing the difficulty of various management departments, through the continuous updating of architectural design models, the completed photogrammetric model will be combined with the BIM indoor model to realize the normal application of architectural design and park management and improve the overall informatization level.

**4.4. Updating strategy for plant space protection.** The renewal of the plant space focuses on the value of protecting industrial heritage. In view of the existing problems in the plant, based on the Iterative reconstruction model of the plant, the space composition, building conditions and environmental performance of the plant are analyzed, and the functions of the park and the leading industries are positioned, the spatial composition of the park has been determined, and the physical space and cultural environment of the park have been planned and designed. The specific strategy is shown in Figure 4.2. The renovation of the park includes material space renovation and cultural space renovation, and the material space includes buildings and their external environment, during the update process, following the basic principles of historical building renovation, protecting the original buildings from damage, maintaining the original architectural style, and preserving and inheriting the original historical and social values [11, 2, 4].

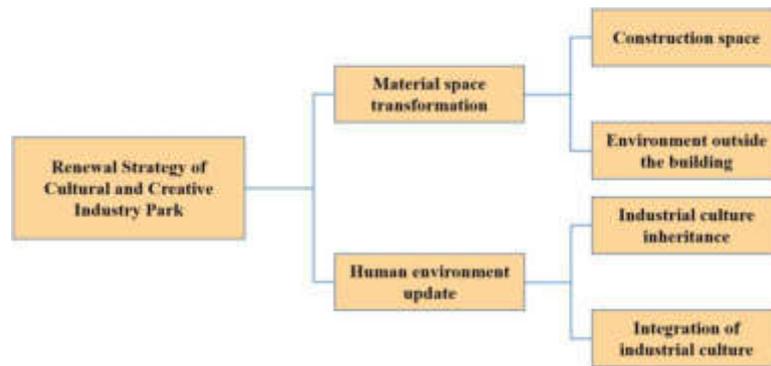


Fig. 4.2: Reconstruction of creative industries park

**5. Conclusion.** With the rapid development of modern society, the protection and utilization of industrial heritage has become a hot issue. The protection and utilization method of industrial heritage based on GA algorithm and Iterative reconstruction proposed by the author aims to protect industrial heritage by means of modern science and technology, by establishing 3D models, optimizing model parameters, and other technical means, the goal of 3D modeling, post model management, and extracting the value of industrial heritage is achieved. Finally, by positioning the protection and renovation of a certain factory area, the goals and principles of protection and renovation are determined, according to the Iterative reconstruction model and field survey, the buildings of the plant are classified and counted, the creative attributes of the park are determined, the spatial structure of the park is adjusted, and the plant is protected and renovated based on the Iterative reconstruction results. In the renovation, music is the theme, integrating new functions with industrial architecture to achieve new forms of expression of old industrial heritage, revitalizing industrial heritage and transforming it into a new driving force for urban development. The GA algorithm was used to optimize the 3D model, thereby improving the accuracy of the data and making it more in line with the actual situation.

## REFERENCES

- [1] H. BENNANI, B. McCANE, AND J. CORNWALL, *Three-dimensional reconstruction of in vivo human lumbar spine from biplanar radiographs*, Computerized Medical Imaging and Graphics, 96 (2022), p. 102011.
- [2] M. GALLO, M. DI NARDO, L. C. SANTILLO, ET AL., *A simulation based approach to support risk assessment*, in Recent Advances in Automatic Control, Modelling and Simulation, ATHENS, WSEAS Press, 2013, pp. 86–92.
- [3] R. GENONI, D. PIZZOCRI, F. ANTONELLO, T. BARANI, L. LUZZI, T. PAVLOV, J. GIGLIO, AND F. CAPPIA, *Three-dimensional reconstruction from experimental two-dimensional images: Application to irradiated metallic fuel*, Journal of Nuclear Materials, 548 (2021), p. 152843.
- [4] ———, *Three-dimensional reconstruction from experimental two-dimensional images: Application to irradiated metallic fuel*, Journal of Nuclear Materials, 548 (2021), p. 152843.
- [5] C. LI, G. DONG, AND R. WANG, *A three-dimensional reconstruction algorithm of nonwoven fabric based on an anthill model*, Textile Research Journal, 92 (2022), pp. 1876–1890.
- [6] Y. LI, J. GAO, X. WANG, Y. CHEN, AND Y. HE, *Depth camera based remote three-dimensional reconstruction using incremental point cloud compression*, Computers and Electrical Engineering, 99 (2022), p. 107767.
- [7] P. LIANG, W. ZHENG, AND X. XU, *Three-dimensional reconstruction method based on jitter optimization*, 1914 (2021), p. 012006.
- [8] S. LYU, Y. WANG, X. LI, J. GUO, C. WANG, AND Q. ZHANG, *Three-dimensional reconstruction of coronal mass ejections by the correlation-aided reconstruction technique through different stereoscopic angles of the solar terrestrial relations observatory twin spacecraft*, The Astrophysical Journal, 909 (2021), p. 182.
- [9] ———, *Three-dimensional reconstruction of coronal mass ejections by the correlation-aided reconstruction technique through different stereoscopic angles of the solar terrestrial relations observatory twin spacecraft*, The Astrophysical Journal, 909 (2021), p. 182.
- [10] Y. PENG, Z. WU, G. CAO, S. WANG, H. WU, C. LIU, AND Z. PENG, *Three-dimensional reconstruction of wear particles by multi-view contour fitting and dense point-cloud interpolation*, Measurement, 181 (2021), p. 109638.
- [11] H. SONG, L. ZHANG, F. CAO, H. SHEN, Z. NING, Y. HUANG, X. ZHAO, X. GU, Z. QIU, AND J. SUN, *Three-dimensional reconstruction of bifilm defects*, Scripta Materialia, 191 (2021), pp. 179–184.

- [12] X. TENG, G. ZHOU, Y. WU, C. HUANG, W. DONG, AND S. XU, *Three-dimensional reconstruction method of rapeseed plants in the whole growth period using rgb-d camera*, *Sensors*, 21 (2021), p. 4628.
- [13] F. ULLAH, S. MIRAN, F. AHMAD, AND I. ULLAH, *A low-cost three-dimensional reconstruction and monitoring system using digital fringe projection*, *Transactions of the Canadian Society for Mechanical Engineering*, 45 (2020), pp. 331–345.
- [14] ———, *A low-cost three-dimensional reconstruction and monitoring system using digital fringe projection*, *Transactions of the Canadian Society for Mechanical Engineering*, 45 (2020), pp. 331–345.
- [15] X. WANG, Z.-Y. YIN, J.-Q. ZHANG, H. XIONG, AND D. SU, *Three-dimensional reconstruction of realistic stone-based materials with controllable stone inclusion geometries*, *Construction and Building Materials*, 305 (2021), p. 124240.
- [16] L. XIN, X. LIU, Z. YANG, X. ZHANG, Z. GAO, AND Z. LIU, *Three-dimensional reconstruction of super-resolved white-light interferograms based on deep learning*, *Optics and Lasers in Engineering*, 145 (2021), p. 106663.
- [17] H. YOO AND J.-Y. JANG, *Computational three-dimensional reconstruction in diffraction grating imaging by convolution with periodic  $\delta$ -function array*, *Optik*, 249 (2022), p. 168211.
- [18] Y. ZHANG, X. JIANG, Z. MA, Q. YANG, Y. ZENG, B. DENG, AND H. WANG, *An efficient interpolation-free algorithm for cylindrical millimeter-wave holographic three-dimensional reconstruction*, *Microwave and Optical Technology Letters*, 63 (2021), pp. 1018–1023.
- [19] J. ZHAO, C. FENG, G. CAI, R. ZHANG, Z. CHEN, Y. CHENG, AND B. XU, *Three-dimensional reconstruction and measurement of fuel assemblies for sodium-cooled fast reactor using linear structured light*, *Annals of Nuclear Energy*, 160 (2021), p. 108397.
- [20] S. ZHENG, R. HE, H. CHEN, Z. WANG, X. HUANG, AND S. LIU, *Three-dimensional reconstruction and sulfate ions transportation of interfacial transition zone in concrete under dry-wet cycles*, *Construction and Building Materials*, 291 (2021), p. 123370.

*Edited by:* B. Nagaraj M.E.

*Special issue on:* Deep Learning-Based Advanced Research Trends in Scalable Computing

*Received:* Aug 25, 2023

*Accepted:* Oct 18, 2023



## LIGHTWEIGHT SALIENCY TARGET INTELLIGENT DETECTION BASED ON MULTI-SCALE FEATURE ADAPTIVE FUSION

MUQING ZHU\*

**Abstract.** In order to solve the problems of small targets, variable shooting angles, and heights in drone images, the author proposes an adaptive drone target intelligent detection algorithm based on multi-scale feature fusion. The results show that after adding a deconvolution cascade structure to the network, mAP increased by about 2.5 percentage points and AP<sup>50</sup> increased by about 3 percentage points. Compared with Method 3, Method 4 uses GA-RPN instead of RPN, and when the IOU is 75, the AP increases by 3.5 percentage points, reflecting that the target prediction candidate boxes generated using semantic features adaptively match better than the manually designed target candidate boxes. This indicates that the proposed target detection framework has better classification ability and higher frame regression accuracy. Multi scale adaptive candidate regions are used to generate fused features of different scales generated by the network, weighted fused multi-scale features are used for target prediction, and semantic features are used to guide the network to adaptively generate target candidate frames, greatly enhancing the feature expression ability of various targets and improving the detection accuracy of aerial targets.

**Key words:** Target detection, Deep network, Feature fusion, Multi-scale feature adaptation, Lightweight residual network model

**1. Introduction.** Over the past two decades, the computer vision field has shifted its emphasis from traditional approaches to an increasing focus on vision-based learning. The evolution of research in this field has seen a notable transition from conventional methods to the prominent use of deep learning techniques. The application of deep learning has gained significant importance in addressing various challenges within the realm of computer vision [12]. Productivity research is an important aspect of computer vision. It is the basis of high resolution and high level of clarity (such as separate images), image processing, scene understanding, object detection, image description, etc. The core task of target detection is to label the target concerned by the task and the target location information in the image to be detected. Since the 21st century, the performance of computing devices has been greatly improved, and the field of target detection relying on high-performance computing devices has also ushered in considerable development [8]. At the same time, the demand for intelligent industrial production continues to expand, and the market demand for target detection technology is also growing, research in the field of target detection has already begun. Currently, in-depth research-based target detection technology has been applied to autonomous driving systems; face recognition; medical images; security; fault diagnosis; military and other field [3]. The object detection method identifies the most attractive targets from the input image and is the initial step of a multi-vision computer [4]. When considering the evolution of saliency target detection, it can be categorized into two distinct approaches: traditional methods that rely on manually crafted features and heuristic priors, and task-oriented saliency target detection methods built upon deep learning. Traditional saliency target detection primarily depends on specific features like color, texture, and image gradients to compute target significance. While these methods are capable of identifying important elements in an image, they are constrained by the need for extensive prior data on significance, which can limit their effectiveness in complex environments. Traditional detection methods have low detection efficiency and long detection time. The saliency target detection algorithm based on deep learning benefits from the rapid development of Full Convolution Network (FCN), and its performance is far better than the traditional methods. FCN has powerful feature extraction ability, which can obtain edge details, texture clues, context features and high-level semantic information with multi-layer and multi-scale. However, as the number of network layers is stacked, pooling operation brings high-level abstract semantic expression, and at the same time, it also leads to image size reduction, thus losing a lot of detail information. In recent years,

---

\*Guangzhou Huali College, Guangzhou, Guangdong, 511325 (Corresponding author, [MuqingZhu9@163.com](mailto:MuqingZhu9@163.com))

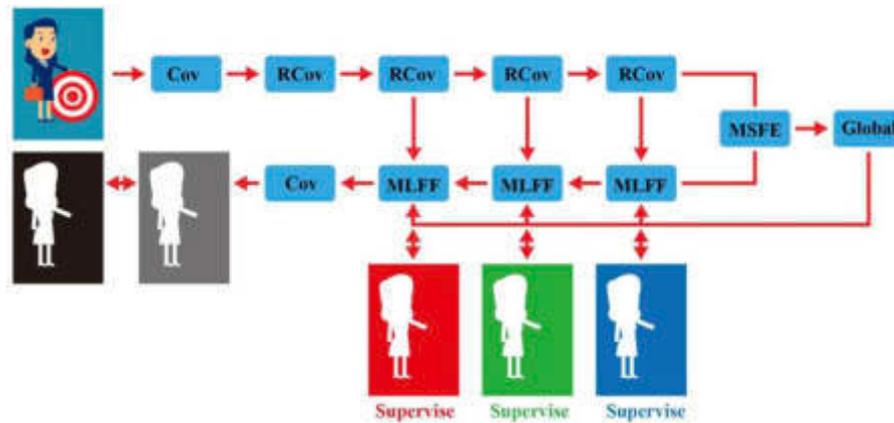


Fig. 1.1: Network framework of multi-scale feature extraction and multi-level feature fusion

a large number of end-to-end saliency detection frameworks have a large number of parameter redundancy in the encoding and decoding stages, resulting in a decline in the test speed, which cannot meet the real-time requirements. Convolution neural network can extract different levels of features, among which the advanced features have semantic features and can locate significant objects; The lower level features have rich details and can be used to sharpen the edges of prominent objects. A direct idea is to simply splice different levels of features. However, it is difficult to achieve the desired effect and capture clearer details in this way. In order to solve the above problems, in order to improve the recognition accuracy of small and medium-sized targets in low performance UAV aerial imagery, this paper presents a modified UAV The goal is to find problems based on multiple combinations. As shown in Figure 1.1, the network framework for multi-scale feature extraction and multi-level feature fusion [11].

**2. Adaptive UAV Target Detection Algorithm Based on Multi-scale Feature Fusion.** To enhance the precision of detecting small objects in less efficient aerial images captured by UAVs, we've introduced an adaptive algorithm for UAV object detection. This algorithm is grounded in the fusion of multiple features, ensuring improved accuracy in recognizing these small targets [5]. The first part is the Light Source Deep network (LResnet) for content extraction. Combining the advantages of residual learning, the ordinary solving tasks are divided into the deep network processing and the fast solution process, which improves the efficiency of the network. The second part is a multi-scale adaptive candidate area generation network, select the last layer C2, C3, C4, C5 of the same size output feature map generated by each layer from the four layers of LResnet, and leverage  $1 \times 1$  The length of each solution is fixed at 256 by the solution. The cascade deconvolution model is employed to enhance the resolution of deep-layer feature maps. This, in turn, enables the utilization of the broader contextual information from upper-layer feature maps. Weighted fusion is applied to these maps, taking into account their channel sizes. This process results in the extraction of four distinct feature points, denoted as P2, P3, P4, P5, which exhibit robust orientation cues for target detection [6]. On each layer of features generated by deconvolution cascade network, GA-RPN (Guided Anchoring Region ProposalNetwork) is used to adaptively generate candidate boxes and corresponding category probability values for prediction targets according to semantic features, and the final prediction results are obtained through non maximum suppression [2].

**2.1. Lightweight residual network model.** Deeper degree and smaller receiver area of solving neural network can improve the accuracy of distribution network. The traditional solution function is to add channel parameters and convolution kernel of the input process after the solution, and the output is used as the content of the next layer. However, with the increase of the network depth, the traditional algorithms and numerical algorithms increase with the depth of the network layer, which leads to the increase of the model size, which is difficult to implement for UAV platforms with limited resources. In order to solve this problem, the depth

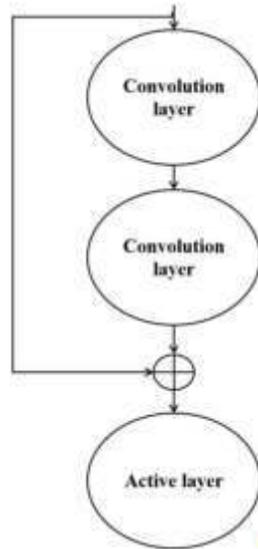


Fig. 2.1: Schematic diagram of residual structure

separation algorithm was used to optimize the model solution, that is, the model solution was decomposed into the depth and the algorithm. In the depth resolution, each channel of the input image is provided with a resolution kernel, and each resolution kernel is only responsible for the resolution of the channel image. Then use  $1 \times 1$ , compress and combine the output results of each channel obtained through depth convolution [10].

After the ordinary convolution is decomposed into depth convolution and point convolution, although the calculation of network model parameters is effectively reduced, the number of network layers is greatly deepened, and the gradient is easy to disappear during network training, leading to greater difficulty in model training, that is, “network regression” phenomenon. The use of residual structure can greatly reduce the burden of deep network training, connecting the shallow network with the deep network through SkipConnection is equivalent to fusing the underlying feature information into the high-level, which can ensure that the input feature information is not lost, and enhance the expression ability of features to the target, and the gradient can be well transferred to the shallow layer. In conclusion, a lightweight depth residual network model (LResnet) is constructed to extract the convolution features of aerial images. In the network structure parameters of the model, OutputSize represents the output feature size, Kernel represents the convolution kernel size, and OutputChannels represents the dimension of the output feature [7].

The residual network (ResNet) is a new deep learning network, which can reach 152 layers of convolution at the deepest. To address the issue of overfitting in deep neural networks, the author introduced a “residual” structure designed to mitigate the problem of vanishing gradients. This structure is illustrated in Figure 2.1.

The activation function is to solve the problem that the linear model in the neural network is not strong enough to add nonlinear factors. Through this function, the features are retained, the redundancy in some data is removed, and finally mapped out. The activation functions of neural network include linear function, threshold function, sigmoid function, bipolar S-shape function, hyperbolic tangent function, and now commonly used ReLU (corrected linear units) function. Apply certain mathematical and physical principles to achieve the desired effect. The network activation function used in this paper is the nonlinear PReLU function (equation (2.1)). Parameters of the negative part in PReLU  $\alpha$  It is learnable rather than fixed. It is only necessary to give him an initial value during training, and then it can be continuously revised according to the depth of training. Compared with ReLU, PReLU introduces additional parameters, so there is no need to worry about over-fitting, and the convergence speed is faster.

$$f(x) = \begin{cases} x & x > 0 \\ \alpha(e^x - 1) & x \leq 0 \end{cases} \quad (2.1)$$

Table 2.1: Parameters of each deconvolution layer

Layer	Type	Kernel	Stride	Output size
h1	Deconvolution	3×3	1	14×14×256
h2	Deconvolution	3×3	1	28×28×256
h3	Deconvolution	3×3	1	56×56×256

**2.2. Multi scale adaptive candidate area generation network.** The aerial image of UAV is very large, which leads to the small proportion of targets in aerial image and the insufficient analysis content. When using depth convolution neural network to extract the image's objective features, the resolution neural network has high acceptance space and strong data representation ability, but geometric data representation ability of weak features, which is not good for small target detection; The fundamental network has a compact size, excelling at capturing geometric information while falling short in representing semantic information effectively. In addition, the traditional target detection network uses manually designed fixed size candidate boxes, different size candidate boxes need to be designed for different detection problems, and the size of the target varies greatly, the fixed design will hinder the improvement of detection accuracy. In order to solve the above problems, a multi-scale adaptive candidate region generation network is constructed based on the LResnet framework. High level semantic features are weighted into the low level feature map through deconvolution cascade structure. It enhances the expression ability of features on the target, and uses multi-level and different scale feature maps for target prediction, on each layer of features generated by the deconvolution cascade network, the position and shape of candidate frames are predicted according to image features, and sparse and arbitrarily shaped candidate frames are generated. Parameters of each layer of deconvolution cascade structure are shown in Table 2.1.

**2.3. Multi-task loss function.** Because the candidate region is generated adaptively, the anchor location loss function  $L_{loc}$  and anchor shape loss function are added on the basis of the traditional classification loss and regression loss. The total target loss function can be expressed as formula (2.2)

$$L = L_{cls} + L_{reg} + \beta_1 L_{loc} + \beta_2 L_{shape} \quad (2.2)$$

where  $\beta_1$  and  $\beta_2$  are the weighting coefficients of the multi-task loss function, with values of 1 and 0.1 respectively.

**3. Results and Analysis.** The experimental platform adopts i7-7700 processor, NVIDIA GTX1080Ti graphics card, 16G memory, and Ubuntu 16.04 operating system [9]. The experimental data used by the author is from the target detection data set of VisDrone UAV, including urban, rural, park, road and other natural scenes, it is obtained by the UAV platform in different positions and at different heights. There are 10 predefined categories marked in the data set, namely pedestrians, people, cars, trucks, buses, trucks, motorcycles, bicycles, tricycles with sheds and tricycles (where "pedestrians" refer to people with standing or walking postures, and "people" refer to people with other postures). Because the flight altitude and camera direction of the UAV are constantly changing, most of the target scales and shooting angles in this data set change greatly, and small targets account for a large proportion, in some data, targets are densely distributed.

**3.1. Qualitative result analysis of target detection.** In order to verify the effectiveness of the proposed solution in scheduling operations, the actual application scene image that is difficult to detect the UAV aerial target was extracted for testing, and the observed results of the algorithm are analyzed and shown. The algorithm in this paper can detect most targets. The results show that the algorithm uses multi-level competition between candidates in the electric field to make multi-level integration of low-level and high-level features, and use the elements semantic language to guide the network to create competitive plans, which improves the detection accuracy of small-scale aerial targets. When faced with a large number of objects and occlusion, the algorithm can also have good detection results, which shows that the algorithm in this paper uses various features to predict objects, and greatly improves the ability of feature extraction. In addition, the algorithm is very little affected by light changes, and it can still have good detection performance in the dark

Table 3.1: Comparison of feature extraction networks

Model	Size /MB	Ratio /%	Accuracy /%
Resnet	97.7	-	81.3
L Resnet	10.2	10.4	80.6

Table 3.2: Effectiveness comparison test of algorithm modules

Method	mAP	AP <sup>50</sup>	AP <sup>75</sup>
1Faster-RCN( Resnet50+ RPN)	18.63	35.87	17.86
2L Resnet+RPN	18.52	35.75	17.44
3LResnet+DC+ RPN	21.03	38.46	18.03
4LResnet+DC+ GA-R PN( ours)	22.1 2	38.76	21.53

environment illuminated by night lights, this also proves that the proposed algorithm can cope with a variety of environmental changes, meet the needs of actual tasks, and has a good generalization ability.

**3.2. Algorithm feasibility verification analysis.** To evaluate the feature extraction network’s performance, an experiment was conducted to compare LResNet with ResNet. The experimental data validated the feature extraction network’s efficacy when employed by the proposed algorithm. To quickly assess the algorithm’s performance, a training set comprising 13,700 images from the VOC2012 dataset was selected, along with a test set of 3,425 images, all under the same conditions. The results, presented in Table 3.1, highlight that LResNet’s network model is remarkably compact, with a size of only 10.2 MB, which is roughly one-tenth the size of ResNet-50’s network model. However, under the same conditions, the classification accuracy of LResNet and ResNet on VOC2012 is only 0.7% different, in addition, the video memory occupation rate required by the LResNet model in the running phase is also greatly reduced, only 546 MB of video memory is needed, which is about 20% of the ResNet-50 network video memory occupation rate. This shows that LResNet greatly reduces the amount of network parameters while losing very little detection accuracy, it also greatly reduces the memory usage during algorithm running.

In order to verify the effectiveness of deconvolution cascade model (hereinafter referred to as DC module) and GA-RPN in UAV aerial target detection algorithms in this algorithm, a multi-scale comparison scheme was designed based on VisDrone target detection data, as shown in Table 3.2 and Figure 3.1. The test experiment was conducted in the VisDrone test dev dataset (1610 UAV aerial images, including various situations in the VisDrone dataset), set Faster Rcn (Resnet50+RPN) as the baseline comparison network, and conduct quantitative analysis using the evaluation indicators of mean precision (mAP) and average precision (AP, including APs with IOUs of 0.50 and 0.75, recorded as AP<sup>50</sup> and AP<sup>75</sup>) [1].

By comparing Method 1 and Method 2 in Table 3.2, it can be seen that after the algorithm replaces the Resnet50 with a large number of parameters with a lightweight residual network, compared with Faster RCNN, the mAP decreases by only 0.11 percentage points, it shows that the network model optimized by depth separable convolution has little impact on the detection effect, but it greatly reduces the number of network parameters [13]. By comparing Method 2 and Method 3, it can be seen that after deconvolution cascade model is added into the network, the mAP increases by about 2.5 percent content, and the AP<sup>50</sup> reaches about 3 percent content. This shows that the algorithm in this paper combines the weights in the low-level map into the high-level sequence map by using the cascade decision model, and the multi-level combination algorithm gets a more robust scheme. It uses multi-level information to predict the target. It is more suitable for multi-target weather forecast that changes with UAV flight altitude. Compared with method 3, method 4 replaces RPN with GA-RPN, when IOU is 75, the AP increases by 3.5 percentage points, which reflects that the candidate frame for target prediction generated adaptively using semantic features is more matched than the target candidate frame designed artificially, it also indicates that the target detection frame proposed by the author has better classification ability and higher frame regression accuracy.

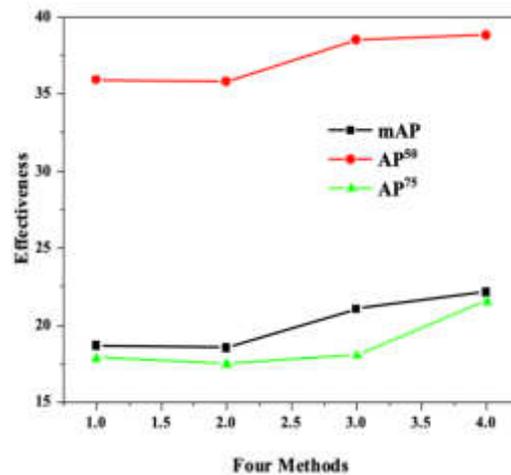


Fig. 3.1: Comparison of effectiveness of each module of algorithm

Table 3.3: Comparison of test results of various categories in VisDrone test data set

Method	Pedestrian	Person	Bicycle	Car	Van	Truck	Tricyele	Awn	Bus	Motor
Faster-RCNN	18.34	7.62	6.76	43.31	27.53	19.95	10.13	7.65	36.87	8.79
Ours	22.43	7.61	8.56	50.18	34.63	24.34	14.1	9.08	36.25	14.88

The detection results of the algorithm for each category in the Vis Drone test dataset are shown in Table 3.3, the comparison index is the AP of various targets. As can be seen from the Table, because Faster RCNN algorithm uses single depth feature and manually designed candidate box to predict targets, it can not adapt well to the actual situation of variable target scales and many small targets in UAV aerial photography data set, so it has poor detection effect for most types of targets. Compared with Faster RCNN, the APs of other categories except “people” and “buses” proposed by the author have improved by 1~7 percentage points. Among them, the average detection accuracy of “pedestrian”, “bicycle” and “motorcycle”, which occupy a small proportion in the image, has been significantly improved, indicating that the proposed algorithm uses multi-scale adaptive candidate regions to generate fusion features of different scales generated by the network, at the same time, multiple scale features after weighted fusion are used for target prediction, and semantic features are used to guide the network to adaptively generate target candidate boxes, which greatly enhances the feature expression ability of various targets and improves the detection accuracy of aerial targets [14].

**3.3. Comparison experiment of mainstream UAV target detection algorithms.** During the comparative experiment phase, we assess models derived from various popular target detection networks after subjecting them to the same training data, mainly comparing the ability of different target detection algorithms to detect and recognize ground targets when the UAV is flying at low altitude, and verifying the detection performance of this algorithm. The comparison algorithm includes RetinaNet, FPN, YOLOv3 and CornerNet. In the experiment, the evaluation indexes of mAP, AP<sup>50</sup>, AP<sup>75</sup> and frame rate (FPS) were used to quantitatively analyze the detection accuracy and detection speed.

The results displayed in Figure 3.2 and Figure 3.3 demonstrate a significant enhancement in the accuracy of target detection indicators compared to mainstream target detection algorithms when applied to UAV aerial photography data. The mean Average Precision (mAP) has notably reached 22.12%, surpassing YOLOv3 by 1.82%. The AP<sup>50</sup> serves as an effective metric to evaluate the algorithm’s classification performance, while AP<sup>75</sup> reflects the ability of the detection framework to precisely determine bounding box positions. Comparative experiments reveal that the algorithm proposed in this paper achieves a detection accuracy of 21.53% at an

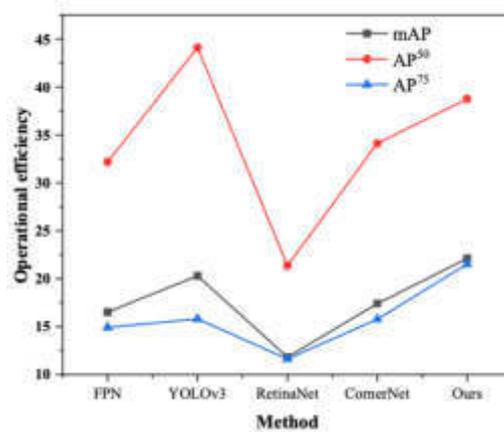


Fig. 3.2: Comparison of UAV aerial photography data of mainstream target detection algorithm (mAP, AP<sup>50</sup>, AP<sup>75</sup>)

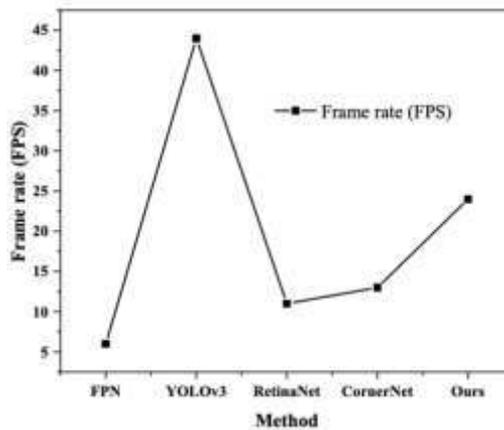


Fig. 3.3: Comparison of UAV aerial photography frame rate (FPS) of mainstream target detection algorithm

Intersection over Union (IOU) of 0.75, representing a substantial 5.73% improvement over YOLOv3. This underscores the superior classification capability and enhanced accuracy in bounding box regression of the target detection framework introduced in this paper. The notable improvement in detection accuracy can be attributed to the multi-scale adaptive candidate region generation network, which assigns weight to high-level semantic features within the low-level feature map through the deconvolution cascade structure. This approach leverages multi-level, multi-scale feature maps for target prediction, significantly enhancing the detection of diverse target types influenced by changes in UAV perspective and flight altitude. Moreover, for each feature layer produced by the deconvolution cascade network, candidate frame positions and shapes are predicted based on image features. This results in the generation of sparse and arbitrarily shaped candidate frames, which aren't constrained by fixed sizes, and therefore, they align more closely with actual target frames. This directly contributes to improved accuracy in frame regression. The detection speed of the algorithm in this paper has been significantly improved compared with the two-stage target detection algorithms FPN, RetinaNet, etc. Although it has not reached the speed level of YOLOv3, it still has the detection speed of  $24 \text{ frame} \cdot \text{s}^{-1}$ . The main reason is that LResnet has greatly reduced the number of network parameters and improved the operation efficiency of the algorithm.

**4. Conclusion.** Due to the swift advancements in science and technology, unmanned aerial vehicles (UAVs) have found extensive utility across both military and civilian sectors. UAV aerial targeting systems have become important topics in artificial intelligence and computer vision. Focusing on the problems of small measurement, high contrast and high resolution of targets in UAV aerial images, a modified UAV target detection problem is many combinations. Based on the quality of the depth separation algorithm and the residual learning, a light extraction network is created. A large number of candidate adaptive region generating network are designed, and the maps with different spatial coefficients are weighted and fused according to the channel dimensions, which improve the expression ability of the target characteristics. The key points are used to guide the network to adapt to the target parameters which match the real target better of multi- target map. The experiments showed that the algorithm successfully improves the detection accuracy of the UAV target air and has good power.

## REFERENCES

- [1] C. BAOYUAN, L. YITONG, AND S. KUN, *Research on object detection method based on ff-yolo for complex scenes*, IEEE Access, 9 (2021), pp. 127950–127960.
- [2] M. GONG AND Y. SHU, *Real-time detection and motion recognition of human moving objects based on deep learning and multi-scale feature fusion in video*, IEEE Access, 8 (2020), pp. 25811–25822.
- [3] Z. HUO, X. LI, Y. QIAO, P. ZHOU, AND J. WANG, *Efficient photorealistic style transfer with multi-order image statistics*, Applied Intelligence, 52 (2022), pp. 12533–12545.
- [4] P. JIA AND F. LIU, *Lightweight feature enhancement network for single-shot object detection*, Sensors, 21 (2021), p. 1066.
- [5] B. LI AND Y. HE, *A feature-extraction-based lightweight convolutional and recurrent neural networks adaptive computing model for container terminal liner handling volume forecasting*, Discrete Dynamics in Nature and Society, 2021 (2021), pp. 1–17.
- [6] G. LI, X. HAO, L. ZHA, AND A. CHEN, *An outstanding adaptive multi-feature fusion yolov3 algorithm for the small target detection in remote sensing images*, Pattern Analysis and Applications, 25 (2022), pp. 951–962.
- [7] Y. LI, S. ZHANG, AND W.-Q. WANG, *A lightweight faster r-cnn for ship detection in sar images*, IEEE Geoscience and Remote Sensing Letters, 19 (2020), pp. 1–5.
- [8] F. QINGYUN, Z. LIN, AND W. ZHAOKUI, *An efficient feature pyramid network for object detection in remote sensing imagery*, IEEE Access, 8 (2020), pp. 93058–93068.
- [9] Q. RAN, Q. WANG, B. ZHAO, Y. WU, S. PU, AND Z. LI, *Lightweight oriented object detection using multiscale context and enhanced channel attention in remote sensing images*, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 14 (2021), pp. 5786–5795.
- [10] Y. SHI, J. LI, Y. ZHENG, B. XI, AND Y. LI, *Hyperspectral target detection with roi feature transformation and multiscale spectral attention*, IEEE Transactions on Geoscience and Remote Sensing, 59 (2020), pp. 5071–5084.
- [11] L. WANG, L. XU, J. SHI, J. SHEN, AND F. HUANG, *Lightweight adaptive enhanced attention network for image super-resolution*, Multimedia Tools and Applications, 81 (2022), pp. 6513–6537.
- [12] L. YANYU AND L. JINBAO, *Small objects detection method based on multi-scale non-local attention network*, Journal of Frontiers of Computer Science & Technology, 14 (2020), p. 1744.
- [13] J. ZHANG, Y. MENG, AND Z. CHEN, *A small target detection method based on deep learning with considerate feature and effectively expanded sample size*, IEEE Access, 9 (2021), pp. 96559–96572.
- [14] M. ZHANG, Y. CHEN, X. LIU, B. LV, AND J. WANG, *Adaptive anchor networks for multi-scale object detection in remote sensing images*, IEEE Access, 8 (2020), pp. 57552–57565.

*Edited by:* B. Nagaraj M.E.

*Special issue on:* Deep Learning-Based Advanced Research Trends in Scalable Computing

*Received:* Aug 25, 2023

*Accepted:* Oct 18, 2023



## OPTIMIZATION OF RADIO ENERGY TRANSMISSION SYSTEM EFFICIENCY BASED ON GENETIC ALGORITHM

RUIJUAN DU\*

**Abstract.** In order to better understand the efficiency optimization problem of radio energy transmission systems, the author proposes a genetic algorithm based research on efficiency optimization of radio energy transmission systems. The author first addresses the issue of improving the efficiency of magnetic coupled radio energy transmission. On the basis of ensuring a certain transmission distance and voltage gain, the system is mathematically modeled using coupling circuit theory, and mathematical expressions such as transmission efficiency, transmission distance, and voltage gain are obtained as the objective functions of the algorithm. Secondly, the impact of metal obstacles on the transmission system was analyzed. Design a radio energy transmission compensation circuit, and through simulation, obtain three transmission system parameter schemes that meet the objective function and constraint conditions. Finally, the multi-objective genetic algorithm is used to optimize the system parameter design and obtain the optimal combination of transmission system parameters, with coupling coefficient  $k=0.1818$  and mutual inductance coefficient  $M=23.165 \times 10^{-5}H$ . Using multi-objective genetic algorithm, the algorithm has a fast convergence function in terms of the number of iterations, a non dominated function solution, and Pareto graphs have verified that the numerical value (3) in the text is the optimal combination design for the transmission system.

**Key words:** Genetic algorithm, radio energy transmission, electromagnetic coupling

**1. Introduction.** The wireless energy transmission network is a planar self-organizing network architecture, where the nodes within the network are wireless power supply devices. For example, sensors in wireless sensor networks and robots in robot football matches can be regarded as nodes of a wireless energy transmission network. Each node has three functions, namely energy transmission, energy reception, and energy relay. For limited nodes distributed within a certain spatial range, electrical energy is emitted from the transmitting node and transmitted one-on-one through several relay nodes to reach the receiving node. A power transmission link is formed between the transmitting node and the receiving node, several transmission links form the network topology of a wireless energy transmission network. Link transmission efficiency: Electric energy passes from the transmitting node to the receiving node through the relay node, and the ratio of the electric energy picked up by the receiving node coil to the electric energy emitted by the transmitting node coil is the transmission efficiency of the link. The efficiency between nodes is related to the mutual inductance between coils. The larger the mutual inductance, the stronger the coupling between coils, and the higher the transmission efficiency. However, the mutual inductance between coils is related to distance, and the larger the distance, the smaller the mutual inductance. Scholars have derived a formula for calculating the mutual inductance between any coil. The author calculates the mutual inductance between nodes based on this formula, and then calculates the efficiency between nodes. The transmission efficiency of adjacent two nodes is multiplied sequentially to obtain the transmission efficiency of the link. Link lifetime: The link of WPTNs dynamically changes based on the electrical energy requirements of nodes in the current network. The period from the formation of the transmission link to the end of energy transmission is called the link lifetime. Self repair ability: The ability of a link to self repair and restore the transmission of electrical energy when a malfunction occurs during transmission and the link cannot function properly. Non crossability: In WPTNs, nodes transmit electricity in a one-to-one transmission mode, and a node can only work in one state. There can be multiple transmission links at the same time, but the transmission links do not cross each other. Threshold electrical energy: Nodes can only participate in the power transmission of WPTNs when their stored energy is above the threshold electrical energy.

When WPTNs establish a link, the receiving node sends an energy request signal outward, and adjacent

---

\*School of Intelligent Manufacturing, Nanyang Institute of Technology, Nanyang, Henan 473000, China (Corresponding author).

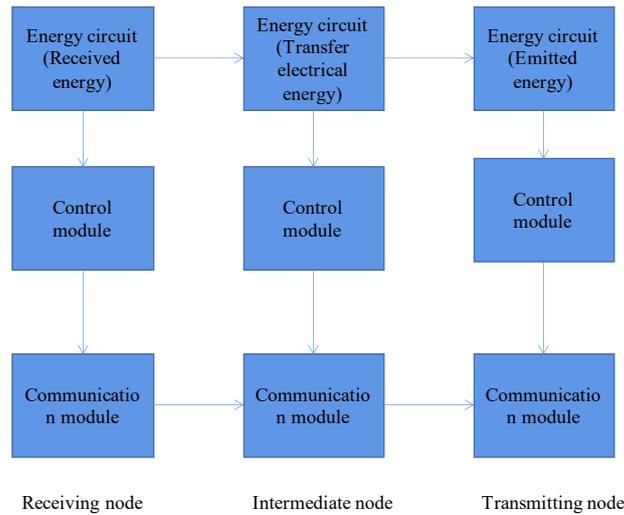


Fig. 1.1: 3 Node Transmission Link.

nodes receive the energy request signal. Based on its own working state and energy storage situation, the receiving node responds by receiving signals from neighboring nodes and selecting one of the many feasible nodes to provide energy. The node consists of three parts: energy circuit, control module, and communication module. The energy circuit has three functions: receiving electrical energy, transmitting electrical energy, and transmitting electrical energy. This means that the node has three working states, and can switch between these three states. However, the node only has one function when working, namely receiving electrical energy, transmitting electrical energy, or transmitting electrical energy. As shown in Figure 1.1, this is a radio energy transmission link composed of three nodes. The control module of the receiving node gives corresponding instructions to the communication module, which sends electricity demand information to surrounding nodes. The communication modules of the relay node and the transmitting node upload the information to their respective control modules. After confirming the connection between the receiving node and the relay node, they respectively control their energy circuits to work in the receiving and transmitting electricity states. The power transmission link between the three nodes is thus established [17, 2].

**2. Literature review.** In recent years, wireless power transfer (WPT) technology has undergone rapid development. In terms of rescue and medical applications, wireless energy transmission technology is a key technology for robots to achieve continuous, lightweight, and wireless cable limited work in special environments (such as underwater, metal pipelines, human bodies, etc.). There are four main principles of wireless transmission, such as electromagnetic radiation, ultrasonic, magnetic field coupling, and electric field coupling. The magnetic coupling resonance method ensures transmission distance, transmission power, and transmission efficiency in special environments, with relatively small radiation to the surrounding area.

The application of wireless energy transmission technology in specific occasions such as electric vehicle charging, power supply for built-in devices in the human body, and power supply for high-voltage devices is more convenient, safe, and reliable compared to traditional contact based energy transmission technology. Because there is no direct cable connection between its power supply and electrical equipment, there will be no skin breakage or wear. With the continuous development and application of wireless energy transmission technology, this technology is receiving increasing attention from domestic and foreign researchers. Among them, magnetic coupling resonant radio energy transmission technology is a new type of radio energy transmission technology that has been widely applied in recent years, it breaks the contradiction between the transmission efficiency and distance of traditional wireless energy transmission technology. The parameter optimization of the magnetic coupling system is crucial for improving the transmission efficiency of the system. Electromagnetic coupling

resonant radio energy transmission utilizes non radiative electromagnetic fields in the near field area to complete electrical energy transmission. On the one hand, compared to electromagnetic induction energy transmission, it has greatly improved the transmission distance, bringing greater freedom for electrical equipment to obtain electrical energy; On the other hand, compared to electromagnetic radiation energy transfer, the energy in the near field area has the characteristic of non radiation, so this technology has good safety, and the design difficulty of the electric energy transmitter and receiver is also reduced. In recent years, WPT has become a research hotspot in domestic and foreign research institutions. Some studies have shown that when the transmission distance is reduced to a certain level, the system is in an overcoupled state. In this state, the system exhibits two 'resonant frequencies', and the power received by the load near these two frequency points is high. This phenomenon indicates that the frequency of the resonant system undergoes splitting in an overcoupled state. Due to the phenomenon of frequency splitting, if the system cannot automatically adjust and track the system's operating frequency, the operating frequency and resonant frequency will definitely be inconsistent, and the system will operate in a non resonant state. The power received by the load shedding will inevitably be very low, which will affect the normal operation of the load when it is lower than the rated power. Therefore, monitoring and controlling the operating frequency of the system is an effective means to improve the load receiving power and efficiency.

Tabataba, P. K. M. et al. proposed through circuit theory modeling analysis that the resonance frequency will shift due to factors such as circuit and environment in radio energy transmission. When in an overcoupled state during transmission, frequency splitting occurs, which affects the consistency of the highest efficiency point and power point, and elucidates the characteristics of radio energy transmission systems [13]. LingzhaoWANG et al. constructed a steady-state circuit model of a wireless transmission system considering the influence of higher-order harmonics, and obtained mathematical formulas for system circuit parameters and state variables through mathematical modeling. By using genetic algorithms, the problem of identifying system circuit parameters is transformed into an optimization problem [15]. Devi, R. R., and others proposed using an improved genetic algorithm to optimize the resonant frequency, coupling coefficient, coil turns, coil radius, capacitance, mutual inductance, and other factors of the transmission system, achieving an energy efficiency of 87 for the capsule robot wireless energy transmission system prototype 6 mV, with a transmission efficiency of 8.02% [5].

The author uses the theory of coupled circuits to analyze the circuit model of wireless energy transmission systems and derives a mathematical formula for the voltage gain coefficient of the coupled circuit. By combining the formula of transmission distance and mutual inductance, a transmission and reception circuit is designed through MATLAB simulation. Under the requirements of transmission efficiency, voltage gain, and coupling coefficient, various parameters of the transmission system circuit are obtained. The multi-objective genetic algorithm was used to optimize the parameters of the radio energy transmission system, verifying the effectiveness of the simulation system parameter design [9, 4].

### 3. Methods.

**3.1. Adaptive frequency tracking WPT system.** High frequency power supply is a key issue in magnetic coupling resonant transmission systems, including high-frequency inverter, dual transistor Class E high-frequency resonant, and microsignal power amplification. Although high-frequency inverter has high power and high frequency, its work efficiency is low and its frequency adjustability is poor. The micro signal power amplification system has received more attention and research due to its high work efficiency, easy frequency adjustment, and mature implementation methods. The adaptive frequency tracking WPT system is shown in Figure 3.1.

Using direct digital synthesis (DDS) to generate the source frequency, compared to high-frequency inverter power supplies, it has higher efficiency, higher frequency, and better adjustability [1]. This system uses DSP to control DDS to generate 3.6-5.6 MHz signals as the system's RF signal source. The DDS frequency can reach 20 MHz, and the swing rate of general power operational amplifiers is not enough, resulting in distortion of the output waveform. Therefore, a two-stage current mode amplifier circuit composed of XST13050 is adopted, with a bandwidth of 425MHz, a maximum output current of 0.2 A, and a swing rate of 6.6 kV/us, meeting the design requirements of this system. After wireless transmission, the energy is converted into DC by the AC/DC module, and then transmitted to the load. The load end uses a power detection module to detect the power (Zigbee is selected as the wireless signal transmission device), and then the power and power variables

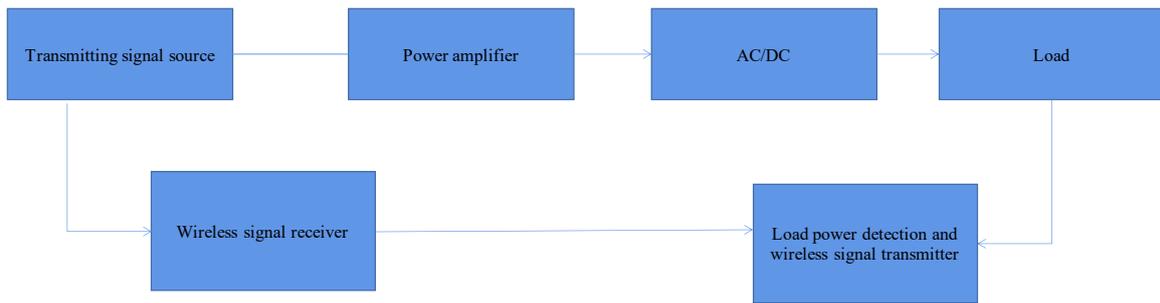


Fig. 3.1: Adaptive frequency tracking WPT system.

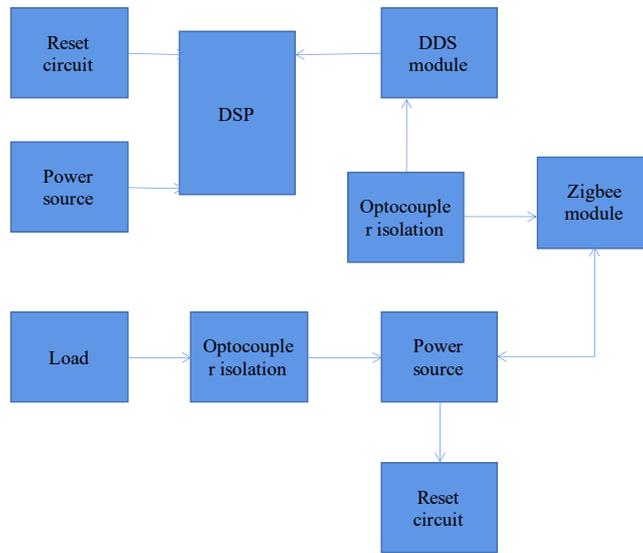


Fig. 3.2: Zigbee hardware circuit.

are transmitted to the Zigbee receiving end through the Zigbee transmitting end, and then to the DSP (the Zigbee receiving end is controlled by the DSP), the DSP immediately triggers an improved genetic algorithm for iterative calculation (calculating the system operation once and generating the corresponding power value), and takes the frequency generated under convergence conditions after 27 iterations as the ideal source frequency for this distance [3].

The wireless communication device in this system adopts Zigbee, with a working frequency of 2550 MHz, a baud rate of 9500, a working channel of Channel20, and a transmission distance of 4 meters. The Zigbee communication module consists of two boards, the transmitter and receiver. The transmitting end and power detection module are both controlled by the load end MCU, and the receiving end is controlled by DSP. The specific hardware circuit is shown in Figure 3.2.

**3.2. Modeling and analysis of transmission system.** The common resonant radio energy transmission system mainly consists of the following three parts: Power module, resonator module, and load module. The power module is used to generate high-frequency current, thereby driving the transmission coil in the resonator module. The resonator module consists of a transmitting coil, a compensating capacitor at the transmitting end, a receiving coil, and a compensating capacitor at the receiving end. The load module end is composed of high-frequency rectifier and voltage regulator, and the load forms a fork. The high-frequency power supply provides high-frequency current to the transmitting coil through a high-frequency inverter circuit. Due to the

Table 3.1: Setting parameters for each parameter in the simulation coupling circuit diagram.

Parameter	Value (1)	Value (2)	Value (3)
Power and load resistance	11	29.8	7.3
$R_{source}; R_{Load}/\Omega$	0	0	0
Primary inductance L1/H	$0.93 \times 10^{-5}$	$310 \times 10^{-5}$	$255.08 \times 10^{-5}$
Secondary inductance L2/H	$0.93 \times 10^{-5}$	$43.18 \times 10^{-5}$	$52.637 \times 10^{-5}$
Primary and secondary capacitance $C_1, C_2/F$	$245 \times 10^{-8}$	$245 \times 10^{-8}$	$245 \times 10^{-8}$
Primary and secondary resistance $R_1, R_2/\Omega$	0.2	0.2	0.2
Coupling coefficients $K_1, K_2$	0.4	0.2007	0.1818
Mutual inductance coefficient M/H	$22.834 \times 10^{-5}$	$22.834 \times 10^{-5}$	$23.165 \times 10^{-5}$
Frequency $f_0/Hz$	$2 \times 10^{12}$	$52 \times 10^4$	$46 \times 10^4$

resonance frequency being equal to that of the receiving coil, the circuit in the transmitting coil generates resonance, and its current generates an electromagnetic field of the same frequency that is coupled to the receiving coil by the near field. It is then supplied to the load through a rectification circuit to achieve wireless energy transmission [10, 18].

According to the equivalent circuit and KVL law, the circuit equation is listed as follows (3.1):

$$\begin{bmatrix} R_0 + j\omega L_0 + j\omega M_{10} \\ j\omega M_{10} R_1 + j\omega L_1 + \frac{1}{j\omega C_1} \end{bmatrix} \quad (3.1)$$

According to equation (3.1), the output voltage gain of the two transmission coil system becomes equation (3.2):

$$s_{22} = \frac{\omega k_{12} \sqrt{L_1, L_2, R_l}}{Z'_1 (Z'_2 + R_L) + \omega^2 k_{12}^3 L_1 L_2} \quad (3.2)$$

In the equation,  $Z$  represents the equivalent impedance of the entire circuit. When the system is in a resonant state, equation (3.3) can be expressed as:

$$s_{22} = \frac{\omega k_{12} \sqrt{L_1, L_2, R_l}}{R_1 (R_2 + R_L) + \omega^2 k_{12}^3 L_1 L_2} \quad (3.3)$$

**3.3. MATLAB simulation data processing.** Simulate using MATLAB software Simulink. Under the premise of ensuring transmission efficiency and voltage gain in the simulation diagram, the parameter settings are shown in Table 3.1.

Provide three wireless transmission system design parameters based on simulation experimental values (1)~(3). The system voltage gain must not be less than 85%. The system coupling coefficient  $K$  is 0.3, and the design parameters should be as close as possible to ensure that the transmission system has a relatively long transmission distance. At the same time, resistors 289, 7.3, and 12  $\Omega$  are used as load resistors for three schemes, and the three design objectives are transformed into multi-objective optimization problem models. Based on the multi-objective genetic algorithm proposed above, the three optimization problem models are solved to design a transmission system scheme that meets the objective function. Firstly, set the variation range of each circuit parameter in the system, assuming that the relative variation range of load resistance is [-1.6%, 1.4%]. The relative variation range of other circuit parameters is [-3%, 3%]. Set the population to 110 and the number of iterations to 220 [7, 16]. The multi-objective genetic optimization algorithm can be used to search for the optimal objective function and corresponding optimal circuit parameter scheme that meets the design objectives. The specific objective function is as follows (3.4)-(3.8).

1) Objective function 1:

$$s_{22} = \frac{\omega k_{12} \sqrt{L_1, L_2, R_l}}{R_1 (R_2 + R_L) + \omega^2 k_{12}^3 L_1 L_2} \quad (3.4)$$

$$0.8 < S'_{22} = \frac{put}{pin} < 1$$

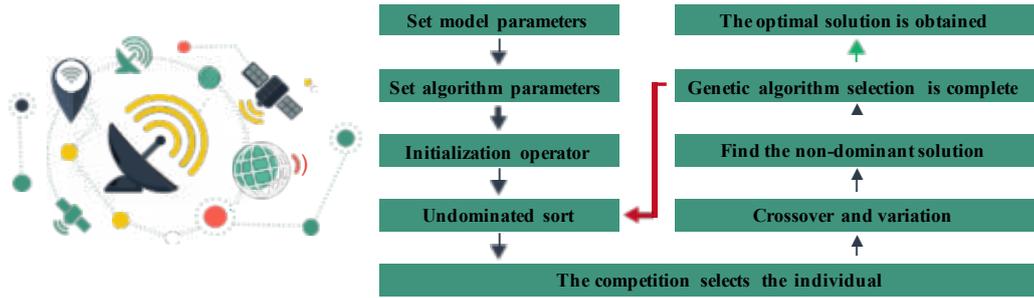


Fig. 4.1: Genetic algorithm process.

2) Objective function 2:

$$M = \frac{\pi \mu n r^4}{2D^3} \tag{3.5}$$

3) Coupling coefficient:

$$k = \frac{M}{\sqrt{L_2 L_1}} < 0.4 \tag{3.6}$$

4) Load resistance:

$$R_L(1 - 0.016) \leq R_L(1 + 0.016) \tag{3.7}$$

5) Resistance of primary and secondary circuits:

$$R_2(1 - 0.4) \leq R_1 = R_2 = 0.1\Omega \leq R_2(1 + 0.4) \tag{3.8}$$

#### 4. Experiments.

**4.1. Validation of data validity based on genetic algorithms.** Using genetic algorithms to simulate the process of biological evolution through computer simulation for search and evolution, and ultimately seeking the optimal solution. The parameter design and implementation process of the magnetic coupling WPT system based on genetic algorithm proposed by the author is shown in Figure 4.1 [12, 11].

#### 4.2. Algorithm process.

1) Set model parameters, set the number of objective functions and variables; Set the upper and lower limits of model parameters based on constraint conditions.

2) Set the algorithm parameters, including the number of chromosomes included in the solution set (population) to 320, the maximum number of iterations to 210, the mutation rate to 0.2%, and the crossover rate to 0.9%.

3) According to the encoding method, initialize the chromosomes and generate the initial chromosomes completely randomly. The initial group data should generally not be too large or too small, and 25-210 is generally selected, which is conducive to the stable convergence speed of the search [20].

4) Fast non dominated sorting, using non dominated sorting algorithms to layer populations of size n. The set of non dominated individuals is the first level non dominated layer of the population. Then, ignoring these marked non dominated individuals and following the aforementioned hierarchical steps, the first level non dominated layer is obtained. And so on, until the entire population is stratified. The fast non dominated sorting of genetic algorithms reduces computational complexity.

5) The tournament method selects individuals, and the tournament selection strategy selects a certain number of individuals from the population each time, and then selects the best one to enter the offspring

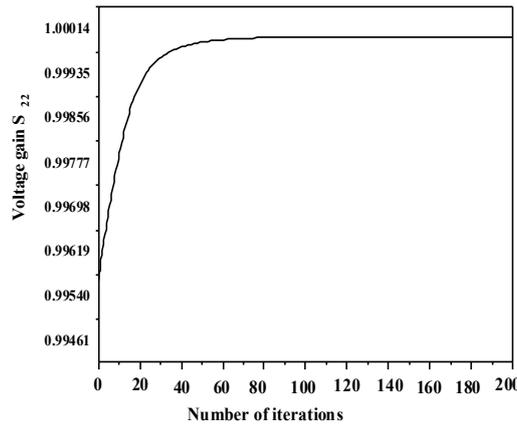


Fig. 4.2: Iterative curve of genetic optimization for system voltage gain coefficient.

population. The competition selection method in genetic algorithm is a return sampling. Repeat this operation until the new population size reaches the original population size. A few yuan tournament is a process of extracting several individuals from the population at once, and then selecting the optimal individual from these individuals and placing them in a set reserved for the next generation of population. Repeat this operation several times as many individuals need to be saved.

6) By using single point mutation and encoding with numerical values and text, methods such as swapping genes at any position with their preceding or following genes or other genes at any position can be used [6, 14].  $r_1 = 2$ , then the variation of chromosomes is equation (4.1):

$$[0.2, 0.6, 1.4, 0, 0.5] \rightarrow [0.2, 0.3, 1.4, 0.5] \tag{4.1}$$

7) Crossing, using two point crossing, first determines the crossing position of two parental chromosomes. The operation of forming a child chromosome from a part of the paternal chromosome and a part of the other paternal chromosome obtains two offspring chromosomes, also known as single point crossing.

Selected two paternal chromosome formulas (4.2):

$$[0.2, 0.6, 1.4, 0, 0.5][0.4, 0.23, 0.43, 0.9, 0.8]_{(10)} \tag{4.2}$$

$$r_1 = 2, r_2 = 5$$

So the cross process is equation (4.3):

$$[02, 06, 1.4, 0, 05][04, 023, Q43, 09, Q9] \tag{4.3}$$

Cross Posterior (4.4):

$$[0.2, 0.23, 0.43, 0.9, 0.5][0.4, 0.6, 1.4, 0.8] \tag{4.4}$$

8) Find non repetitive non dominant solutions and combine the population of the previous generation with the population of the current generation after mutation crossover into one population. Adopting an elite selection strategy, it merges the parent population with the offspring population, allowing the next generation's population to be selected from twice the space, thereby preserving all the best individuals and ensuring that certain excellent population individuals are not discarded during the evolution process.

**4.3. Numerical simulation analysis.** The numerical simulation analysis is shown in Figures 4.2-4.4.

From the analysis of Figures 4.2-4.4, it can be seen that using the values in Table 3.1 (3) for system parameter design meets the optimal requirements for both the voltage gain coefficient and transmission distance objective

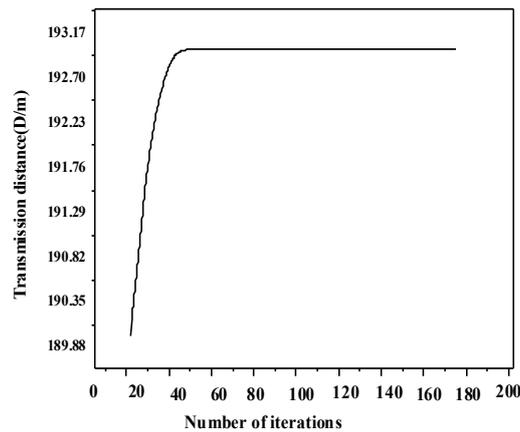


Fig. 4.3: Optimization iterative curve of transmission distance objective function by genetic algorithm.

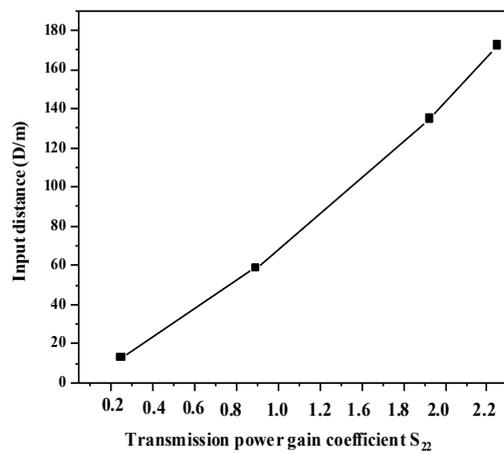


Fig. 4.4: Pareto solution of genetic algorithm optimization objective function.

function of the transmission system. As shown in Figures 4.2 and 4.3, genetic algorithm is used to obtain the iterative curve of the objective function with good convergence, fast convergence speed, and strong search ability. This indicates that the algorithm is effective. The two objective functions in Figure 4.4 are completely in the same direction and have only one non dominant solution [8, 19].

**5. Conclusion.** The author derived a mathematical formula for the voltage gain of a radio energy transmission system through the theoretical knowledge of coupled circuits, and analyzed the impact of eddy current effect generated when metal obstacles exist on the transmission efficiency of the transmission system. That is, the eddy current effect generated by good conductor obstacles reacts with the impedance of the transmitting and receiving coils, increasing the impedance and thereby affecting the voltage gain coefficient of the transmission system. Then, a circuit model of the wireless energy transmission system was built through MATLAB simulation, and three sets of system parameter values were obtained. When the objective function condition is met: The voltage gain coefficient is not less than 0.9, the transmission distance is ensured as far as possible, and the coupling coefficient is not higher than 0.4. There are also transmission system constraints. Using multi-objective genetic algorithm, the algorithm's iteration function converges quickly, there is a non dominated function solution, and Pareto graph validates that the value (3) in Table 3.1 is the optimal combination

design for the transmission system.

## REFERENCES

- [1] W. G. BIN, Z. Y. FAN, W. CHAO, W. X. CHUANG, M. JUN, L. XIAO-DONG, AND Z. LI, *Multi-level peer-to-peer collaborative optimization of smart energy system based on big data analysis*, in Journal of Physics: Conference Series, vol. 1965, IOP Publishing, 2021, p. 012145.
- [2] I. S. BULUT AND H. ILHAN, *Energy harvesting optimization of uplink-noma system for iot networks based on channel capacity analysis using the water cycle algorithm*, IEEE Transactions on Green Communications and Networking, 5 (2020), pp. 291–307.
- [3] J. CUI, H. LI, D. ZHANG, Y. XU, AND F. XIE, *Multi-objective optimization of hydro-viscous flexible drive for dynamic characteristics using genetic algorithm*, Industrial Lubrication and Tribology, 73 (2021), pp. 1003–1010.
- [4] T. DELEBARRE, V. GRAS, F. MAUCONDUIT, A. VIGNAUD, N. BOULANT, AND L. CIOBANU, *Efficient optimization of chemical exchange saturation transfer mri at 7 t using universal pulses and virtual observation points*, Magnetic Resonance in Medicine, 90 (2023), pp. 51–63.
- [5] R. R. DEVI AND T. SETHUKARASI, *An energy-efficient routing based on a hybrid improved whale artificial ecosystem optimization algorithm in wsn*, Concurrency and Computation: Practice and Experience, 34 (2022), p. e6639.
- [6] G. LIU, Y. QIN, M. LI, Y. LIU, J. ZHENG, X. LIU, AND Y. YIN, *Performance analysis and optimization of a pemfc-caorc system based on 3d construction method of thermodynamic cycle*, Energy Conversion and Management, 247 (2021), p. 114730.
- [7] J. LIU, F. ZHANG, C. CATTAND, H. YANG, AND S. WANQING, *Optimization of atp system based on quantum secure communication and its tracking control strategy*, Shock and Vibration, 2021 (2021), pp. 1–13.
- [8] A. MAHROOGHI AND E. LAKZIAN, *Optimization of wells turbine performance using a hybrid artificial neural fuzzy inference system (anfis)-genetic algorithm (ga)*, Ocean Engineering, 226 (2021), p. 108861.
- [9] R. K. PRASAD AND T. JAYA, *Intelligent spectrum sharing and sensing in cognitive radio network by using aroa (adaptive rider optimization algorithm)*, International Journal of Computational Intelligence and Applications, 22 (2023), p. 2341007.
- [10] S. SAHRAEIAN AND F. EMAMI, *Design and optimization of frequency-tunable terahertz absorbers based on graphene using genetic algorithm*, Optical Engineering, 61 (2022), pp. 067105–067105.
- [11] K. SHEELAVANT AND R. SUMATHI, *Correlative dynamic mapping based optimization (cdmo) for optimal allocation in cognitive radio sensor network*, Revista Gesto Inovao e Tecnologias, 11 (2021), pp. 3955–3973.
- [12] P. SHI, Y. QIN, Y. CAO, H. ZHAO, R. GAO, AND S. LIU, *Topology optimization of metamaterial microstructure for wireless power transfer with high power transmission efficiency*, Journal of Magnetism and Magnetic Materials, 537 (2021), p. 168228.
- [13] F. S. TABATABA, P. ROUHANI, AND M. KOOLIVAND, *Energy efficiency optimization for amplify and forward relay networks with channel estimation errors*, Telecommunication Systems, 76 (2021), pp. 541–552.
- [14] C. TANG, B. XUE, AND L.-X. WANG, *Optimization design of shaped charge based on improved genetic algorithm*, in IOP Conference Series: Materials Science and Engineering, vol. 1043, IOP Publishing, 2021, p. 042034.
- [15] L. WANG AND R. CHOU, *Energy efficiency research of decode-and-forward two-way simultaneous wireless information and power transfer full-duplex relay system*, Journal of Signal Processing, 38 (2022), pp. 806–815.
- [16] B. WU AND S. ZHANG, *Energy management strategy for dual-motor two-speed transmission electric vehicles based on dynamic programming algorithm optimization*, SAE International Journal of Electrified Vehicles, 10 (2021), pp. 19–32.
- [17] G. XIA, J. CHEN, X. TANG, L. ZHAO, AND B. SUN, *Shift quality optimization control of power shift transmission based on particle swarm optimization–genetic algorithm*, Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering, 236 (2022), pp. 872–892.
- [18] Z. YAN, D. YIN, L. CHEN, AND W. SHEN, *Research on the torsional vibration performance of a cvt powertrain with a dual-mass flywheel damper system*, Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering, 236 (2022), pp. 1144–1154.
- [19] S. ZHANG, Y. ZHENG, AND G. LI, *Research on distribution center layout optimization based on genetic algorithm*, in Journal of Physics: Conference Series, vol. 1976, IOP Publishing, 2021, p. 012010.
- [20] G. ZHOU, R. LIU, Z. ZHANG, C. YANG, AND H. DING, *Optimization of diesel engine dual-variable geometry turbocharger regulated two-stage turbocharging system based on radial basis function neural network-quantum genetic algorithm*, Energy Sources, Part A: Recovery, Utilization, and Environmental Effects, (2021), pp. 1–17.

*Edited by:* B. Nagaraj M.E.

*Special issue on:* Deep Learning-Based Advanced Research Trends in Scalable Computing

*Received:* Aug 31, 2023

*Accepted:* Nov 4, 2023



## INTELLIGENT DETECTION AND ANALYSIS OF SOFTWARE VULNERABILITIES BASED ON ENCRYPTION ALGORITHMS AND FEATURE EXTRACTION

HENG LI\*, XINQIANG LI† AND HONGCHANG WEI‡

**Abstract.** Implement status detection of ship software, identify the source of faults in problematic software, and release new software versions. Based on the above requirements, the author regards the detection and control of ship software status as the core research content. Based on the actual operating environment of ship software, the functional requirements of software status detection were studied and analyzed, and a set of ship software status detection was designed and implemented, a software inspection and maintenance platform that integrates ship software operation and maintenance, as well as ship software version release and update. The author conducted practical verification of the SM3 and SM2 hybrid encryption algorithm and selected software on the ship for detection. After analyzing the experimental results, it has been proven that using a hybrid algorithm for encryption and decryption, the server can accurately obtain software information on the ship's platform, detect the software status on the ship, and locate specific problem files. For software that does not meet the standard status, the server can accurately transmit software information to the "component integration framework" and put the component in a "prohibited" scheduling state. After the server repairs the problematic software, the detection results of the software change and display as legal, while the software is in the "allowed" scheduling state in the "component integration framework".

**Key words:** Encryption algorithm, feature extraction, software vulnerabilities, intelligent detection

**1. Introduction.** With the development of integrated electronic warfare equipment, the proportion of software in equipment is becoming larger and more complex. Large software is usually jointly developed by multiple development units responsible for research and development. This cross unit joint research and development strategy improves the efficiency of large-scale software development. On the other hand, the state of system software becomes difficult to control, and the various software cannot cooperate effectively. Maintenance work after failures is also difficult to carry out. After the delivery of equipment software, unauthorized changes to the software are also difficult to control, and local software changes may lead to unstable operation of the entire system software, affecting the normal use of military equipment [9]. Therefore, there is an urgent need to carry out research on software status detection and maintenance related technologies, achieve monitoring and management of ship software status changes through technical means, solve the problem of uncontrolled ship software status changes, and ensure software quality. A software inspection and maintenance system based on component scheduling, firstly, it can solve the problem of uncontrolled status of ship software, allowing users to clearly observe the status of all controlled ship software on each ship station and discover a list of unauthorized software status changes; Secondly, for problematic software, this system can quickly locate the source of the fault, promptly identify and handle the problem, avoiding maintenance personnel from aimlessly debugging and consuming a large amount of time and cost, enabling the problematic software to quickly resume normal operation, at the same time, it also saves a lot of manpower and resources; Finally, this system provides the only legal software release and update interface, and all ship software is deployed to various ship stations through this system, ensuring the unity of the overall software status of the fleet, standardizing the software release and update methods, simplifying the operation process, and enabling developers to focus more on improving the quality and functionality of ship software. In summary, a software detection and maintenance system based on component scheduling provides an effective solution for the detection and comprehensive management of ship software, which is very necessary. In response to this research issue, Yu, L. and others have applied neural net-

---

\*Software Engineering Department, Shijiazhuang Information Engineering Vocational College, Shijiazhuang, Hebei 050000, China

†Mechanical and Electrical Engineering Department, Shijiazhuang Information Engineering Vocational College, Shijiazhuang, Hebei, 050000, China

‡Software Engineering Department, Shijiazhuang Information Engineering Vocational College, Shijiazhuang, Hebei, 050000, China (Corresponding author)

work technology to binary similarity detection, which has become a promising research topic, and vulnerability detection is an important application field of binary similarity detection. Embedding binary code into matrices using neural networks also requires addressing feature representation issues in vulnerability detection. However, current research mostly extracts the syntax or structural features of binary codes, and uses basic blocks as the minimum analysis unit, which is relatively rough. In addition, the structural characteristics of binary functions are usually represented by dependency graphs. During the embedding process, only the neighbor information of the node is obtained, ignoring the global information of the graph. In order to address these two issues, the author proposes a dual channel feature extraction method to obtain finer grained semantic features and globally represent structural features rather than locally [18]. Duan, L. et al. in intelligent systems, attackers can use botnets to launch different network attack activities against the Internet of Things. Traditional botnet detection methods usually use machine learning algorithms, but due to the imbalance of traffic data in the network, it is difficult to detect and control botnets. The author proposes a novel and efficient botnet detection method based on the cooperation between an autoencoder neural network and a decision tree on a given network. Firstly, the deep flow detection method and statistical analysis are used as feature selection techniques to select relevant features for characterizing the communication related behavior between network nodes. Then, the self encoder neural network is used for feature selection to improve the efficiency of model construction. Finally, the Tomak recursive boundary synthesis minority oversampling technique is used to generate additional minority samples to achieve class balance, and an improved gradient enhanced decision tree algorithm is used to train and establish an anomaly traffic detection model to improve class balance [2].

Based on current research, the main functional objectives of this system are to detect the status of ship software, repair faults, and release and update software versions. Maintain and record software states that do not comply with agreed rules, and prohibit the invocation of that component in the “Component Integration Framework”. Backup and save ship software with stable operation and good performance; Release a new version of the software and deploy it to ship stations, or update the old version of the software. Through a software detection and maintenance system platform based on component scheduling, users can accurately grasp the real-time status of ship software, identify problem software, locate problem sources, and strengthen the monitoring ability of ship software; Improve the efficiency of software version release and update, and save resource costs for software status changes; Accurately grasp the software version distribution of the entire fleet and individual ships, issue reasonable combat instructions based on the software status on different ships, and improve coordination and interaction between ships. This system provides comprehensive support for the stable operation of ship software, provides solutions for the management and control of ship software status, and provides technical support for the combat command system. It is of great significance for the control of fleet software and battlefield combat command.

## 2. Methods.

**2.1. SM2 elliptic curve algorithm.** The SM2 algorithm was released by the National Password Administration on December 17, 2010, the asymmetric cryptographic algorithm based on ECC (Elliptic Curve Cryptography) has the characteristics of requiring less private key bit length for operations, lower system parameter requirements, less storage space, lower broadband requirements for data transmission, and lower overall power consumption of the algorithm. Therefore, it can be widely applied to devices with relatively small system scales and severely limited resources [11]. The SM2 algorithm mainly includes three parts of applications: Digital signature algorithm, key exchange protocol, and public key encryption algorithm. With the development and progress of international cryptographic technology, the current 1024 bit RSA algorithm has faced serious security issues, and the elliptic curve cryptography algorithm has its algorithm performance advantages. It has been used as a standard for public key cryptography algorithms in many countries and regions. In order to ensure the security of domestic passwords and improve the security factor, the National Cryptography Administration began researching elliptic curve algorithms with independent intellectual property rights in 2001. After extensive research on public key cryptography algorithms with high recognition in the international cryptology community, learning from advanced cryptographic experience abroad, and absorbing the theoretical foundation of existing elliptic curve algorithms, the SM2 algorithm was successfully completed in 2004, And on December 17, 2010, the SM2 algorithm standard was announced [15]. In March 2011, Bank of China released the relevant specifications for its financial IC card, stating that it uses the SM2 elliptic curve algorithm to enhance the

Table 2.1: Comparison of breakthrough time between RSA algorithm and SM2 algorithm.

RSA Key Strength	SM2 Key Strength	Breakthrough time (year)
512	106	104, has been breached
768	132	108, has been breached
1024	160	1011
2048	210	1020

Table 2.2: Performance comparison of RSA and ECC algorithms.

Algorithm	Signature efficiency (times/second)	Verification efficiency (times/second)
1024 bit RSA	2793	51225
2014 bit RSA	456	15123
256 bit SM2	4096	872

security of IC card applications. At the same time, non-financial applications using PBOC3.0 as a reference standard also use the SM2 algorithm. Comparing and analyzing the two algorithms, SM2 and RSA, it was found that SM2 has higher security than RSA. Under the same key strength, the SM2 algorithm is more difficult to crack and takes longer to break; In terms of computational efficiency, the signature efficiency of SM2 is much higher than that of RSA, while the efficiency of verifying signatures is slower than RSA. The specific comparison between the two algorithms is shown in Tables 2.1 and 2.2:

**2.2. SM3 password hash algorithm.** The hash function, also known as the hash function, its functional feature is that it can output fixed length summary information after a series of changes and processing of an unlimited length message, and the resulting summary result is called a hash value [10, 5]. The hash function has unidirectionality, meaning that its encryption process is easily implemented in a computer, while the decryption process is not feasible in a computer. Therefore, the hash function is considered an irreversible encryption algorithm, and the generated ciphertext is treated as the “fingerprint” of the message or data block, serving as the unique identifier of the message to verify the content of the information. The one-way variation and fixed length output characteristics of hash functions make them widely used in fields such as data integrity verification, digital signatures, message authentication codes, and data cryptographic protocols.

The structure of the compression function of the SM3 algorithm is similar to that of the SHA-256 algorithm, but the SM3 algorithm incorporates many new design techniques, including adding 16 steps of all XOR operation, message doubleword intervention, and adding fast avalanche effect P-permutation, which can effectively avoid high probability local collisions and resist various cryptographic analyses [12]. Moreover, the SM3 algorithm reasonably utilizes word addition operations to form a carry plus 4-stage pipeline. Without significantly increasing hardware overhead, it adopts P-permutation to accelerate the avalanche effect of the algorithm and improve computational efficiency. The SM3 password hash algorithm adopts basic operations suitable for 32b microprocessors and 8b smart cards, which has high efficiency and wide applicability for cross platform implementation. The comparison between the SM3 algorithm and the SHA series algorithm is shown in the Table 2.3:

The software on ships is organized and operated by a series of files and components. The detection of the status of ship software is actually the detection of the content of the files and components that make up the software [8, 7]. Therefore, in order to implement this system module, the author proposes to algorithmically process all file contents of the target software, convert the file contents into file feature values through a series of operations, and package other attributes of the file together into network packets. The server compares the file attributes and file feature values of the ship software with the software standard library, and realizes the software status detection module by analyzing the matching results.

Table 2.3: Comparison of SM3 and SHA series algorithms.

algorithm	The length of the output hash value (bits)	Data block length (bits)	Maximum length of input message (bits)	The length of a word (bits)	Number of cycles	The operator used
SHA-0	160	512	264-1	32	80	+,and,or, xor,rotr
SHA-1	160	512	264-1	32	80	+,and,or, xor,rotr
SHA-256	256	512	264-1	32	64	+,and,or, shr ,xor,rotr
SHA-512	512	1024	2128-1	64	80	+,and,or, shr,xor ,rotr
SM3	256	512	264-1	32	64	+,and,or, shr,xor,rotr

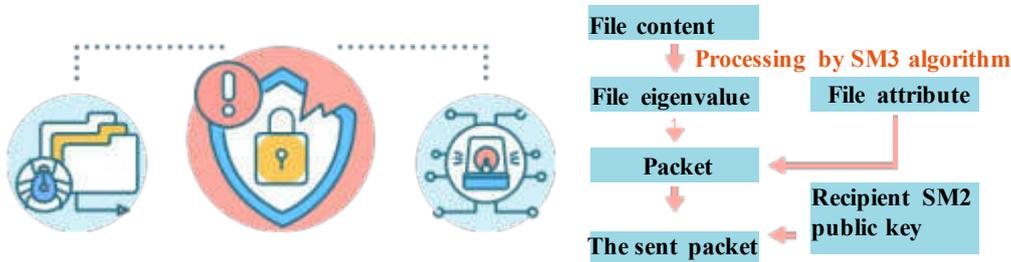


Fig. 2.1: Hybrid algorithm encryption flowchart.

**2.3. Hybrid encryption algorithm technology combining SM3 and SM2.** The author utilizes the advantages of the SM3 algorithm, such as fast computation speed, higher security, lower broadband requirements, and easy key management, in order to combine the two algorithms to obtain a more efficient, accurate, and secure encryption technology [4]. The basic principle of algorithm design is that when the server detects the software status of the ship location, the ship location first uses the SM3 password hash algorithm to process the file content, calculate the file feature values, integrate them with other attributes of the file into a message packet, and then use the SM2 algorithm to encrypt the message packet, the ciphertext containing file information is transmitted from the ship location to the server through a TCP connection. After receiving the ciphertext data sent by the ship location, the server decrypts it using the SM2 decryption algorithm, then compare the obtained file feature values with the feature values in the standard library to determine whether the file's status is legal. This encryption and decryption method not only ensures the security of message data but also improves the speed of message transmission, thereby ensuring the speed, accuracy, and effectiveness of ship software status detection.

(1) Hybrid algorithm encryption process

The encryption process of the hybrid algorithm is to use SM3 to calculate the file feature values based on the file content; The ship station packages the file feature values and other attributes into a message [19]. Encrypt the message using the SM2 public key sent by the server at the ship station; The ship station transmits the encrypted message to the server through a TCP connection. The entire encryption process is shown in Figure 2.1.

(2) Hybrid algorithm decryption process

The decryption process of the hybrid algorithm is as follows: The server decrypts the packet using the SM2 private key to obtain the file attribute information of the ship's location software; Based on the file name, file

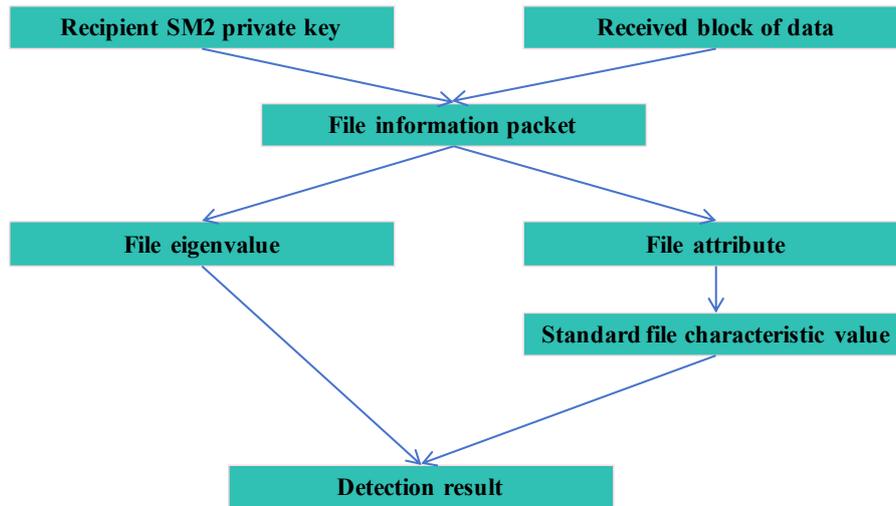


Fig. 2.2: Flow chart of hybrid algorithm decryption.

path, software name, and software version in the file attributes, locate the standard file record corresponding to the file; Compare the feature values of the ship's location file with those in the standard library, determine the status of the file, and display the detection results [16]. The entire decryption process is shown in Figure 2.2.

### 3. System Design.

**3.1. Communication design.** When designing network data communication between the ship's position (client) and the software detection and maintenance system (server), Socket UDP network communication is chosen as the interaction method for the server to issue ship software instructions and report the network status of the ship's position based on the on-site network environment of the system, at the same time, a strategy of triggering high-frequency interaction with a timer is adopted to ensure the speed, real-time performance, and stability of the data; Using SocketTCP as the communication method for software status detection and software entity transmission between servers and ship stations, while adopting flow control and message confirmation strategies to ensure the integrity of data transmission.

**3.2. Architecture design.** The logical architecture of a software inspection and maintenance system based on component scheduling is mainly divided into two levels: Business layer and service layer. The business layer mainly includes visual interfaces for various functions and the basic logic for business implementation, and is the presentation layer for interaction with users. The service layer mainly includes the network data transmission between the server and the ship station, as well as the underlying support of the server's internal database.

**3.3. Deployment design.** The system adopts a CIS structure, with servers deployed in the ship software status command room, and clients distributed on all ship stations that require detection and maintenance [17]. The server has a visual interface and complete functions such as ship and ship software status detection, fault software repair, software version release and update; The client does not have a visual interface and is a console program that starts automatically in the background. It can receive command messages sent by the server and has the function of remote data transmission with the server.

**3.4. Subsystem design.** The software inspection and maintenance system based on component scheduling mainly includes four subsystems: Ship software status detection, software standard library operation and maintenance, software entity control, and software supervision and recording [20]. The main functional objective of the ship software status detection subsystem is to extract the ship software status within the local area network based on the user input of the ship location and relevant information of the ship software. After

comparison with the software standard library, the software status detection results are displayed, mainly including three functional modules: Ship location information control, ship software status detection, and fault analysis of the problematic software. The functional objective of the software standard library operation and maintenance subsystem is mainly to provide a reference template for matching the status of ship software, which is the basis for ship software status detection and fault analysis. It mainly has the function of publishing new software status standards and maintaining software standard status; The main functional objective of the software entity control subsystem is to manage the physical files of the software, mainly including backing up stable and high-performance ship software, repairing ship software with running faults and illegal status, and upgrading and updating outdated ship software. In addition, the system has also designed a subsystem structure for software supervision and recording, with the main functional goal of recording detailed information on the detection results, fault causes, version changes, and other operations of ship software, facilitating users to access and analyze the historical operations of ship software.

## 4. Results and Analysis.

### 4.1. Ship position software detection.

#### (1) Ship position information control

The main function of ship location information control is to manage and control ship location information. The data items stored in the system for ship location include the name of the ship location, the network P of the ship location, and the list of software controlled by the ship location. The software list information includes the name of the software, the version of the software, and the storage path of the software in the ship location system [1]. The ship position information control module should not only provide a visual interface for ship position information, but also provide functions for adding and modifying ship position information. At the same time, for ship fleets using the same software system, a unified software status control list template should be provided to quickly declare new ship position information through the template.

#### (2) Ship software status detection

The server issues detection instructions to the ship location software. The ship location first traverses the file directory of the detected software to obtain the file organization structure of the software. The SM3 algorithm is used to sequentially process the file content, calculate the file feature values, and integrate them with other attributes of the file into a message package. Then, the SM2 algorithm is used to encrypt the message package, transfer the ciphertext containing file information from the ship's location to the server through a TCP connection. After receiving the ciphertext data sent by the ship's location, the server decrypts it using the SM2 decryption algorithm, and then compares the obtained file feature values with the feature values in the standard library to finally determine the status of the file and display it.

#### (3) Problem software fault analysis

Problem software fault analysis mainly focuses on illegal states and ship software that have malfunctioned during operation. Based on the detection of ship software status, fault analysis identifies specific problem sources, identifies files related to the fault and the time of the fault occurrence, enabling users to repair the problem software in a targeted manner, effectively analyzing the cause of the fault, and optimizing the performance of ship software.

**4.2. Software standard library operation and maintenance.** The software standard library operation and maintenance subsystem can release new software status standards, abstract software entities into software feature values, store various data of standard software status in servers, and provide reference basis for software status detection of ship stations [3, 13]. This module can adjust the details of the released software standard status, present the entire software standard status to users in a good interface interaction, facilitate users to find the source of problems with faulty software, and analyze the functional differences between different software versions.

**4.3. Software entity control.** The software entity control subsystem mainly provides functional interfaces for software entity operations, including ship software backup, ship software fault repair, and ship software version upgrade [14]. In order to improve the unity and correlation between software entities and software states, the operation of software entities is closely tied to ship software state detection and software standard library operation and maintenance, this means that each control of software entities corresponds to a

Table 4.1: Detection results of ship software based on hybrid algorithm.

File name	File type	File size	Version Status
1029LBandGlobePara_ sys.xml	xml	2KB	×
1029LHFGlobePara_ sys.xml	xml	3KB	×
1029LHFGlobePara_ _work.xml	xml	3KB	×
Config.xml	xml	2KB	×
Config.xml	xml	2KB	×
Config.xml	xml	2KB	×
DDSRecvLog.txt	txt	:1KB	×
config.ini	ini	1KB	×
1.500000MHz- CW- 2016 08- 10	Wav	1KB	✓
1.txt	txt	26KB	✓
1029HFAutoSearchPara.xml	xml	41KB	✓
1029HFAutoSearchPara_ LF.xml	xml	41KB	✓
1029HFDFTestPara.xml	xml	5KB	✓
1029HFDemoduCtrlPara_ 1.xml	xml	11KB	✓
1029HFDemoduCtrlPara_ 2.xml	xml	11KB	✓
1029HFDemoduCtrlPara_ 3.xml	xml	11KB	✓
1029HFDemoduCtrlPara_ 4.xml	xml	11KB	✓
1029HFDemoduCtrlPara_ 5.xml	xml	11KB	✓
1 029HFDemoduCtrlPara 6.xml	xml	11KB	✓
1029HFDemoduCtrlPara_ 7.xml	xml	41KB	✓

synchronous change in the status of ship software, effectively avoiding the phenomenon of ship software being tampered with and users' uncontrolled management of ship software entities. Moreover, this system provides a unified interface for software entity operations, eliminating the need for users to perform software entity operations on ship platforms. This simplifies the workflow for repairing and upgrading ship software faults, effectively improves the efficiency of software entity control, greatly reduces software maintenance costs, and has constructive significance for the standardization and standardization of software entity operations.

**4.4. Software control records.** The software control record subsystem provides a visual interface for viewing all software control records of users, mainly recording controlled ship information, such as ship location network IP, ship software name and version, operation time, specific operation behavior, and related file statistics, it mainly records the software detection status of ship positions and changes in software versions. The system will record detailed information on the status of illegal software discovered during the detection process and display it to users in a visual list. Users can query specific problem sources by viewing illegal file records, strengthen the management of ship software, and ensure the stability and uniformity of software operation.

**4.5. Test results.** The author conducted practical verification of the SM3 and SM2 hybrid encryption algorithm and selected software on the ship for detection [6]. After analyzing the experimental results, it has been proven that using a hybrid algorithm for encryption and decryption, the server can accurately obtain software information on the ship's platform, detect the software status on the ship, and locate specific problem files. For software that does not meet the standard status, the server can accurately transmit software information to the "component integration framework" and put the component in a "prohibited" scheduling state. After the server repairs the problematic software, the detection results of the software change and display as legal, while the software is in the "allowed" scheduling state in the "component integration framework". The detection results of ship software status are shown in Table 4.1:

**5. Conclusion.** In the software detection and maintenance system designed by the author, ship software status detection is the basic function of the system for subsequent management and maintenance of ship software. In order to ensure the accuracy and uniqueness of the target software detection results, this system traverses the file directory of the target software and uses the SM3 password hash algorithm to calculate the file feature

values. If there are any changes to the file content, the calculated feature values will definitely not be the same, so by matching the feature values in the software standard library, it is possible to determine whether the content of the file has been tampered with, thereby obtaining the detection results of the ship software. And in response to the disadvantage of low security of a single encryption algorithm, this system adopts a mixed encryption scheme of SM3 and SM2 to encrypt the transmission of network messages, ensuring sufficient security of the software detection process and accurate detection results. We have designed a software inspection and maintenance system based on component scheduling, mainly including communication design, deployment design, architecture design, etc, chosen the Qt compiler as the research and development environment for the system, and SQLite as the storage medium for storing software feature values on the server. We have designed a network message structure for data exchange, and implemented a software detection and maintenance system that includes the above solutions. This not only effectively improves the detection efficiency of ship software status, but also ensures the safety of the detection process and the accuracy of the detection results, moreover, it can significantly shorten the repair time of faulty software, save the cost of ship software maintenance, strengthen centralized management of software release and updates, and improve the control ability of ship software.

## REFERENCES

- [1] K. CHEN, L. HU, M. YAO, L. QIAN, AND Y. ZHANG, *Study on intelligent traffic search method based on driver facial feature analysis*, International Journal of Vehicle Information and Communication Systems, 6 (2021), pp. 151–160.
- [2] L. DUAN, J. ZHOU, Y. WU, AND W. XU, *A novel and highly efficient botnet detection algorithm based on network traffic analysis of smart systems*, International Journal of Distributed Sensor Networks, 18 (2022), pp. 182459–182476.
- [3] M. GHARAMANI, A. O’HAGAN, M. ZHOU, AND J. SWEENEY, *Intelligent geodemographic clustering based on neural network and particle swarm optimization*, IEEE Transactions on Systems, Man, and Cybernetics: Systems, 52 (2021), pp. 3746–3756.
- [4] X. HE, Y. YANG, W. ZHOU, W. WANG, P. LIU, AND Y. ZHANG, *Fingerprinting mainstream iot platforms using traffic analysis*, IEEE Internet of Things Journal, 9 (2021), pp. 2083–2093.
- [5] J. JAGANNATHAN AND M. M. PARVEES, *Vulnerability recognition and resurgence in network based on prediction model and cognitive based elucidation*, Journal of Physics: Conference Series, 2070 (2021), p. 012122.
- [6] K. LI, Y. WANG, Y. SHAO, AND X. WU, *Smart boat detection based on feature pyramid network and deformable convolution*, Journal of Physics: Conference Series, 2083 (2021), p. 042018.
- [7] B. LIU, Y. FAN, B. XUE, T. WANG, AND Q. CHAO, *Feature extraction and classification of climate change risks: a bibliometric analysis*, Environmental Monitoring and Assessment, 194 (2022), pp. 1–41.
- [8] Q. MAJEED AND A. FATHI, *A novel method to enhance color spatial feature extraction using evolutionary time-frequency decomposition for presentation-attack detection*, Journal of Ambient Intelligence and Humanized Computing, 14 (2023), pp. 3853–3865.
- [9] H. SHEN, P. FAN, Z. WEI, C. ZHAO, S. ZHOU, AND Q. WU, *Research on transmission equipment defect detection based on edge intelligent analysis*, Journal of Physics: Conference Series, 1828 (2021), p. 012087.
- [10] Z. SHI, A. A. MAMUN, C. KAN, W. TIAN, AND C. LIU, *An lstm-autoencoder based online side channel monitoring approach for cyber-physical attack detection in additive manufacturing*, Journal of Intelligent Manufacturing, 34 (2022), pp. 1815–1831.
- [11] Z. SONG, X. HUANG, C. JI, AND Y. ZHANG, *Intelligent identification method of hydrophobic grade of composite insulator based on efficient ga-yolo former network*, IEEE Transactions on Electrical and Electronic Engineering, 18 (2023), pp. 1160–1175.
- [12] G. TANG, L. YANG, L. ZHANG, W. CAO, L. MENG, H. HE, H. KUANG, F. YANG, AND H. WANG, *An attention-based automatic vulnerability detection approach with ggnn*, International Journal of Machine Learning and Cybernetics, 14 (2023), pp. 3113–3127.
- [13] Z. TONG, C. XING, Z. ZENG, AND B. LAI, *Intelligent tidal lane system based on vehicle attribute recognition algorithm and related laws and regulations*, Journal of Physics: Conference Series, 1972 (2021), p. 012100.
- [14] B. WANG, J. LI, J. LUO, Y. WANG, AND J. GENG, *Intelligent deblending of seismic data based on u-net and transfer learning*, IEEE Transactions on Geoscience and Remote Sensing, 59 (2021), pp. 8885–8894.
- [15] J. XIE, Y. ZHAO, D. ZHU, J. YAN, J. LI, M. QIAO, G. HE, AND S. DENG, *A machine learning-combined flexible sensor for tactile detection and voice recognition*, ACS Applied Materials & Interfaces, 15 (2023), pp. 12551–12559.
- [16] H. XU, M. GUO, N. NEDJAH, J. ZHANG, AND P. LI, *Vehicle and pedestrian detection algorithm based on lightweight yolov3-promote and semi-precision acceleration*, IEEE Transactions on Intelligent Transportation Systems, 23 (2022), pp. 19760–19771.
- [17] H. YILAHUN AND A. HAMDULLA, *Entity extraction based on the combination of information entropy and tf-idf*, International Journal of Reasoning-based Intelligent Systems, 15 (2023), pp. 71–78.
- [18] L. YU, Y. LU, Y. SHEN, H. HUANG, AND K. ZHU, *Bedetector: A two-channel encoding method to detect vulnerabilities based on binary similarity*, IEEE Access, 9 (2021), pp. 51631–51645.
- [19] B. ZHANG AND Z. XI, *A systematic review of binary program vulnerabilities feature extraction and discovery strategy generation*

*methods*, Journal of Physics: Conference Series, 1827 (2021), p. 012090.

- [20] X. ZOU, R. LV, X. LI, H. CHEN, Z. WANG, AND H. YANG, *Intelligent electrical fault detection and recognition based on gray wolf optimization and support vector machine*, Journal of Physics: Conference Series, 2181 (2022), p. 012058.

*Edited by:* B. Nagaraj M.E.

*Special issue on:* Deep Learning-Based Advanced Research Trends in Scalable Computing

*Received:* Aug 31, 2023

*Accepted:* Nov 4, 2023



## APPLICATION OF CONTROL ALGORITHM IN THE DESIGN OF AUTOMATIC CRIMPING DEVICE FOR CONNECTING PIPE AND GROUND WIRE

CONGBING SHENG<sup>\*</sup>; PENG XING<sup>†</sup>; XIUZHONG CAI<sup>‡</sup> AND ZHENG SHAO<sup>§</sup>

**Abstract.** Due to the low effectiveness and poor quality of manual crimping of grounding wires, the author proposes the design of an automatic crimping device for connecting tube grounding wires based on intelligent fully automatic technology. The device consists of a microcontroller, an upper computer control interface, an electric push rod, an infrared sensor, a pressure transmitter, and other devices. The staff used the upper computer monitoring interface to set the relevant parameters for grounding wire crimping, and used X-ray digital imaging technology to measure the crimping size of the grounding wire. The size met the set parameter conditions. Through the PID control algorithm in the microcontroller, the stepper motor was controlled to push the clamp to move, completing the automatic crimping of the grounding wire. The X-ray detection method was introduced to detect the quality of the grounding wire after the crimping was completed. The experimental results show that the average deviation between the measured crimping size of the grounding wire and the actual measurement size by the automatic crimping device is only 0.06 mm, indicating that its measurement results are accurate; The success rate of crimping exceeds 95%. The above experimental results verify that the designed crimping device has high stability and reliability, and good quality detection effect.

**Key words:** Intelligent fully automatic, Crimping of conductor and ground wire, Device design, Single chip microcomputer, PID control algorithm, Radiographic testing methods

**1. Introduction.** During the construction of transmission and transformation lines, the construction of grounding wires is a crucial step, and this technology will play a decisive role in the construction quality of the transmission and transformation lines. During the construction of power transmission and transformation lines, it is necessary to inspect and involve the actual situation to ensure the implementation of the grounding wire project and ensure that the project has a certain degree of standardization. At the same time, through focused research and discussion on the construction technology of grounding wires, the specific application of common grounding wire construction technologies is discussed, providing effective references for promoting innovation and reform of transmission and transformation line construction technology, creating a solid foundation for the development of construction technology, and providing conditions for technological innovation [16]. The use of construction technology for grounding wires can effectively improve the operational efficiency of power lines, ensure the safety of construction during the construction of transmission and transformation lines, and reduce construction risks. For the construction technology of the grounding wire, in order to innovate and improve the technology, it is necessary to improve the technical construction based on the natural and geographical environment of the site during construction, pay attention to the quality of the hydraulic connection of the grounding wire, and create an excellent condition for the stable operation of the transmission and transformation lines [20]. At the same time, construction personnel of power transmission and transformation lines should try to avoid the phenomenon of steel anchors or steel pipe damage, which usually occurs during the crimping process. After hydraulic completion, it is necessary to ensure that the connected pipes are within the same size range, which will increase the construction difficulty of the transmission and transformation lines. However, this can ensure the construction quality of the transmission and transformation lines. The two main connection methods of mechanical clamping and hydraulic connection have played a great role in the construction process, ensuring the smooth progress of construction, not only improving efficiency, but also ensuring the quality of construction [1].

---

<sup>\*</sup>State Grid Henan Electric Power Company, Puyang, Henan,457000, China (Corresponding Author).

<sup>†</sup>State Grid Henan Electric Power Company, Puyang, Henan,457000, China

<sup>‡</sup>State Grid Henan Electric Power Company, Puyang, Henan,457000, China

<sup>§</sup>State Grid Henan Electric Power Company, Puyang, Henan,457000, China

Firstly, the key points of stability. During the construction process of grounding wires, special attention should be paid to the stability of the construction technology. This is because the construction of grounding wires itself belongs to a high standard construction technology, and any link in the construction process is related to the whole. Moreover, construction faults may occur during the actual construction process, and it needs to be judged based on the type of construction technology. Therefore, special attention must be paid to stability during the construction process of the grounding wire project. For example, cracks and deformations are very common in the construction process of hydraulic pliers. This requires construction personnel to adhere to a professional and rigorous work attitude, ensure that all theoretical construction meets the standards and specifications, and strictly review all construction links to ensure the quality of the transmission and transformation lines, in order to ensure the construction quality [8].

Secondly, normative points. In the application process of grounding wire construction technology, it is necessary to be able to fit in with the actual needs of transmission and transformation line engineering. This is because the standardization of grounding wire construction is not fixed, but dynamically changes according to the construction needs of transmission and transformation lines. Therefore, in the construction process of grounding wire engineering, high requirements are put forward for the technical key point of standardization. Generally speaking, regardless of the form of grounding wire construction specifications, the good operation of construction equipment should be ensured, that is, after inspecting and debugging the construction equipment, construction work should be carried out according to the characteristics of the transmission and transformation line construction. After completing the construction work, it is necessary to carefully review the construction results. Once any construction problems are found, timely repairs must be ordered [22].

Thirdly, key points for damage management. During the construction process of power transmission and transformation lines, the grounding wire engineering is easily affected by various factors, resulting in different forms of damage problems. Therefore, during the construction process of the grounding wire project, if any damage problems are found, they must be dealt with according to relevant standards. For example, material deformation, as a common damage problem, if it cannot be dealt with in a timely manner, it may lead to permanent deformation of the material over time, which is extremely detrimental to the smooth progress of transmission and transformation line construction work. Deformation and damage to the steel core aluminum stranded wire are particularly common in the construction of grounding wire engineering. If the damage scale of the steel core aluminum stranded wire has exceeded one fourth of its own, it must be reconnected. If it does not exceed this scale, pressure equipment can be used to correct it.

As a crucial step in the construction process of transmission and transformation lines, the construction of grounding wires has a decisive impact on the quality of construction to some extent. Therefore, in the construction process of transmission and transformation lines, it is necessary to carry out the grounding wire project according to the actual situation, so that the project results are more standardized. In view of this, the following author will focus on exploring the key points of construction technology for grounding wires, and propose the specific applications of common grounding wire construction technologies in the construction of transmission and transformation lines, in order to lay a good foundation for better promoting the development of transmission and transformation line construction.

**2. References.** For ground wire engineering, hydraulic crimping technology is the most common construction technology. Its main principle is to use hydraulic pumps, pressure piers, oil pipes and other equipment to crimp different types of ground wire materials, thus forming a complete transmission line construction project. The biggest advantage of hydraulic crimping technology is its high stability and relatively low probability of safety accidents. But the disadvantage lies in the strong professionalism of this technology, which requires high-level construction personnel to carry out construction operations smoothly. At the same time, the applicability of this technology is relatively narrow, but because it can be applied to modern grounding wire engineering materials, hydraulic crimping technology has a relatively wide range of applications to some extent. At present, in the construction process of hydraulic pressing technology, the construction focus can be mainly divided into four parts: inspection, cutting, cleaning, and positioning. Among them, inspection refers to the quality inspection of wire clamps and hydraulic equipment before the formal implementation of construction technology, in order to ensure the smooth progress of subsequent construction technology; Cutting refers to the wire cutting work carried out on the premise that the hardware foundation is correct. It should be noted that

before carrying out cutting work, priority should be given to confirming the phase and wire type of the wire to ensure that the wire is in a straight state [13]. At the same time, it is strictly prohibited to use oversized pliers and scissors during the cutting process. This is because cutting work often involves cutting to the standard three-quarters position, and the remaining parts need to be manually cut, which requires that the cutting tool should not be too large, otherwise it is easy to have a one-time cutting problem; Cleaning work, as the name suggests, is the cleaning of materials to avoid contamination affecting material quality. It should be noted that gasoline rather than water is required during the cleaning process of the wire material. After cleaning, the material needs to be dried; Positioning mainly involves marking the dimensions on the pressure pipe to form a preliminary positioning for the smooth progress of subsequent crimping work.

The current development of technology has expanded the application scope of intelligent fully automated technology. By utilizing this technology, comprehensive automation of mechanical devices, product processing, and other aspects can be achieved, which not only saves labor, but also improves work efficiency and product quality. Therefore, all industries have entered the era of intelligent fully automated production. Among various control electronic products, single-chip microcontrollers, also known as single chip microcontrollers, are the most powerful. They are composed of multiple logic functional chips integrated together to form a complete system, which can obtain signals from control devices and achieve control of different facilities and equipment through this signal. Microcontrollers have the characteristics of small size, low price, and strong application performance. X-ray digital imaging technology is a new type of non-destructive testing technology that does not cause part loss when detecting parts and has high detection accuracy.

Shindin, E described an effective implementation of a recent simplex type algorithm for precise solutions of discrete continuous linear programming, and compared it with linear programming approximations obtained through time domain discretization for these problems. This implementation overcomes many numerical traps that are often overlooked in theoretical analysis, allowing for better accuracy or acceleration up to several orders of magnitude compared to previous simplex algorithm implementations and state-of-the-art LP solvers that use discretization. Numerical research includes medium-sized, large-scale, and large-scale examples of scheduling problems and fluid processing network control problems. We discussed online and offline optimization settings for various applications and outlined future research directions [11]. Laurini, M consider the finite element approximation of the Bellman equation for optimal control of switched systems. It has been proven that this problem belongs to a special category that we have studied in our previous work, and for this purpose, we have developed an effective solution algorithm. As an application, we propose the problem of generating autonomous vehicle parking strategies in two typical urban parking scenarios. The vehicle is described by four different switching systems, each of which is associated with a penalty term. Through this approach, we obtained a parking path with a small amount of directional changes and a simple structure [5]. The layout design of nuclear power plant pipelines is to find the optimal routing that meets the goals and constraints in the three-dimensional wiring space. However, due to the dense equipment, complex structure, diverse types of pipeline systems, complex layout constraints, and numerous pipelines in nuclear power plants, pipeline layout is a difficult and time-consuming task even for experienced designers. In order to solve the automatic routing problem of pipelines in the three-dimensional wiring space of nuclear power plants, Zhang, J proposed a large space pipeline automatic routing method that combines Dijkstra algorithm with improved algorithm a. Firstly, this method uses the traditional Dijkstra algorithm to identify key vertices in each room of a nuclear power plant, construct a topology routing diagram, and determine the initial channel area of the pipeline. Secondly, divide the space of the layout area into three-dimensional grids, and then identify and preprocess the items in the layout area [19].

Based on the above research background, this article aims to address the quality and efficiency issues in the crimping process of grounding wires. By combining intelligent fully automatic technology and X-ray digital imaging technology, an automatic crimping device for connecting tube grounding wires is designed based on intelligent fully automatic technology, providing a certain basis for improving the level of grounding wire crimping technology.

**3. Design of automatic crimping device for connecting tube and ground wire based on intelligent fully automatic technology.** This article uses a microcontroller as a controller to design an intelligent fully automatic technology for the automatic crimping device of the connecting tube conductor wire. The

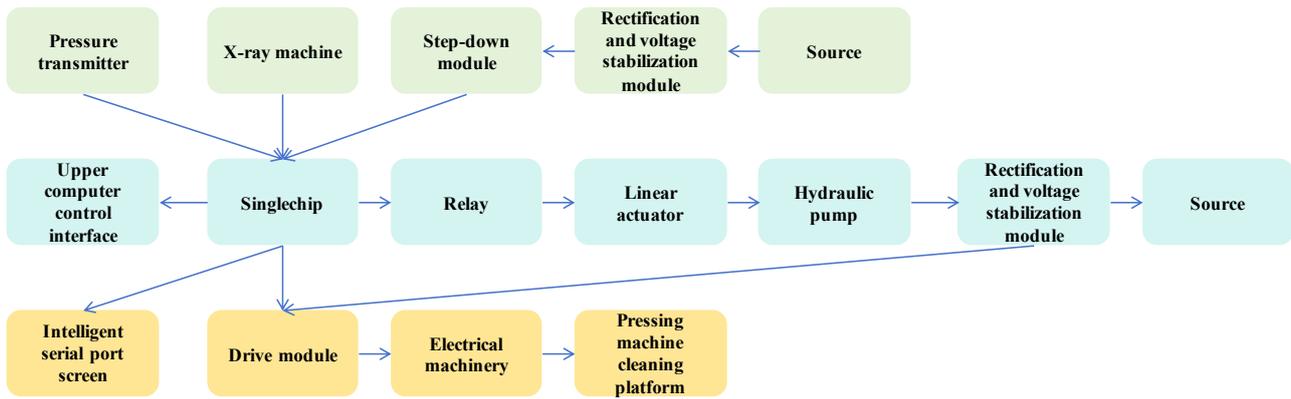


Fig. 3.1: Structure diagram of ground wire automatic crimping device of intelligent automatic technology

automatic crimping device can bear a maximum crimping force of 100MPa, and the measurement accuracy of the conductor wire diameter and movement distance can be accurate to 0.8mm, fully meeting the construction requirements. The automatic crimping device for connecting pipes and ground wires with intelligent fully automatic technology has the advantages of lightweight, simple operation, and ultra-high cost-effectiveness. Figure 3.1 shows the structure of the automatic crimping device for connecting pipes and ground wires using intelligent fully automatic technology [21].

The automatic crimping device for the connecting tube and ground wire of this intelligent fully automatic technology adopts the PCF80C51 series microcontroller from Philips in the Netherlands. This microcontroller has strong data analysis and processing capabilities, can add a large number of breakpoints, and can view memory and registers online, making it easy to operate. The operation process of the microcontroller is as follows: the staff uses the upper computer monitoring interface to set parameters such as the expected diameter and number of movements of the grounding wire after crimping, and transfers the set parameters to the microcontroller controlled by the lower computer controller. The microcontroller controls the stepper motor to push the clamp installed on the pulley to move. When the speed of the stepper motor and the movement of the pulley reach the set parameters, the automatic crimping process of the grounding wire begins. At this time, the changes in the automatic crimping value of the grounding wire are displayed to the staff through the upper computer monitoring interface.

**3.1. Measurement of conductor crimping size based on X-ray digital imaging.** During the automatic crimping process of the grounding wire, there may be some errors in the actual size of the grounding wire crimping due to operational factors by the operator. To avoid the occurrence of crimping errors in the grounding wire, it is necessary to study the correlation between the crimping size of the steel core and the tension of the grounding wire, and accurately measure the actual crimping size of the steel core. After years of development, X-ray digital imaging technology has transformed from simple film radiographic inspection to advanced digital technology, and the detection accuracy is becoming increasingly high. This article uses X-ray digital imaging technology to measure the size of ground wire crimping. This technology consists of digital imaging, image processing, auxiliary facilities, etc. The actual size of the steel core pressed into the steel anchor is obtained by transforming the position and size measurement software of the ground wire, and accurate positioning of the automatic crimping size of the ground wire can be achieved without damaging the ground wire [2]. The specific measurement principle of the crimping size of the grounding wire is shown in Figure 3.2.

### 3.2. Automatic crimping control of conductor and ground wire.

**3.2.1. Crimping pressure control principle.** The pressure of the medium voltage clamp during the automatic crimping process of the grounding wire will directly affect the final effect of the automatic crimping of the grounding wire. Therefore, a pressure detection system is used to accurately measure the pressure of

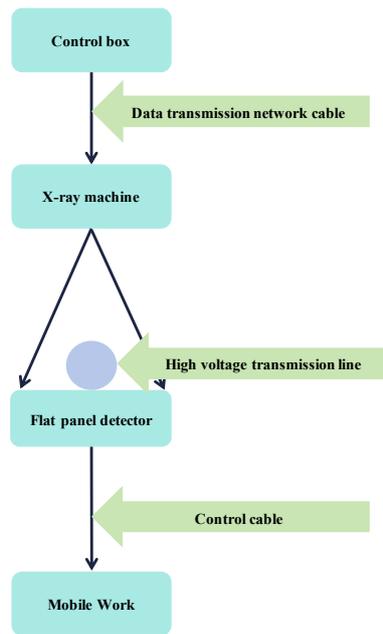


Fig. 3.2: Principle of guide ground ground size measurement of X-ray digital imaging technology

the medium voltage clamp during the crimping process of the grounding wire. The automatic crimping device for the connecting pipe and ground wire of intelligent fully automatic technology outputs liquid oil through the infusion pump driving the crimping pliers. The liquid pressure is detected by a pressure transmitter. To prevent sediment deposition, the pressure tap is set on the side of the process pipeline [7], and the pressure transmitter is set on the edge of the pressure tap.

Set the Wheatstone bridge on the diffusion silicon chip inside the pressure transmitter. When the measured liquid is pressurized, the bridge wall resistance value produces a piezoresistive effect, generating a differential voltage signal [17]. Use an internal dedicated amplifier to convert the signal corresponding to the range into an analog voltage signal, and use the I/O pins of the microcontroller to input the analog voltage signal into the microcontroller. After data processing, output the monitoring results. The automatic crimping device for connecting pipes and ground wires using intelligent fully automatic technology consists of a hydraulic pump, crimping pliers, electric push rods, etc. The hydraulic pump switch is used to control the electric push rod, and the hydraulic pump is connected to the crimping pliers. When the grounding wire reaches the set position, the microcontroller controls the electric push rod to start the hydraulic pump. When the pressure reaches the set value, it starts to press and hold for a certain period of time. Then, the grounding wire is moved to another designated position for a second press, and the above process is repeated to complete the automatic crimping of the grounding wire. The principle of crimping pressure control is shown in Figure 3.3.

**3.2.2. Control algorithm design.** Introducing PID control algorithm in the stepper motor control stage of the microcontroller program can accurately control the movement of the crimped grounding wire. The PID control algorithm can analyze the deviation of the measured value of the ground wire movement and output a certain value to achieve ground wire crimping control [14]. Assuming that the deviation between the actual output value and the theoretical output value is represented by  $u(t)$ , the calculation formula for this deviation is as follows:

$$u(t) = r(t) - b(t) \quad (3.1)$$

Among them,  $r(t)$  and  $b(t)$  respectively represent the set value of the stepper motor speed and the actual measured value of the stepper motor speed.

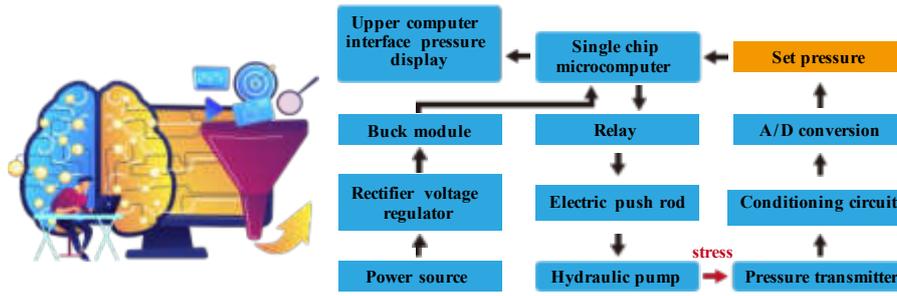


Fig. 3.3: Schematic diagram of the crimp pressure control

Calculate the proportion, integral gain, and differential gain generated by the set value and output value using a linear combination method to achieve microcontroller control. The control calculation formula is as follows:

$$u(t) = K_p \left[ u(t) + \frac{1}{Y_I} \int_0^t u(t)dt + \frac{TDdu(t)}{dt} \right] \tag{3.2}$$

The transfer function form of formula (3.2) after sorting is as follows:

$$G(s) = \frac{U(s)}{E(s)} = K_p \left( 1 + \frac{1}{Y_I s} + Y_D s \right) \tag{3.3}$$

where  $K_p$  represents the proportion;  $Y_I$  represents the integral gain;  $Y_D$  represents differential gain.

To achieve microcontroller control, formula (3.2) needs to be converted into a differential equation, and its calculation formula is as follows:

$$\int_0^t u(t)dt \approx \sum_i^k Y u(i) \tag{3.4}$$

$$\frac{du(t)}{d(t)} \approx \frac{u(k) - u(k - 1)}{Y} \tag{3.5}$$

The PID control calculation formula can be obtained through formulas (3.2), (3.4), and (3.5), as follows:

$$u(k) = K_p \left[ u(k) + \frac{Y}{Y_I} \sum_{i=0}^k e(i) + Y_D \frac{u(k) - u(k - 1)}{Y} \right] \tag{3.6}$$

**3.3. Quality inspection of grounding wire crimping.** During the automatic crimping of the grounding wire, there may be wire drops, resulting in quality issues in the crimping of the grounding wire. This article uses a radiographic testing method to detect wire drops during the automatic crimping process of the grounding wire. The radiographic testing method can clearly identify the crimping status of the grounding wire through image processing, and will not cause loss of the grounding wire during the testing process, saving costs [6].

When a ray passes through a conductive wire, the intensity of the ray is closely related to the density and thickness of the conductive wire. The formula for calculating the attenuation law of radiation is as follows:

$$P = P_0 e^{-\mu x} \tag{3.7}$$

Among them,  $P$ ,  $P_0$ ,  $x$ , and  $\mu$  respectively represent the intensity of the ray after passing through the ground wire, the initial intensity of the incoming ray, the thickness of the ground wire, and the attenuation coefficient of the object.

Table 4.1: Specific parameters of the X-ray machine

Form	Category	Parameter
5*X-ray generator	Radiation source voltage	60~160 kV
	Ray source current	0.2~3 mA
	Focus size	0.4 mm
	2*Maximum penetration force	AL: 100 mm PE:20 mm
4*Imaging detector	Radiation source power	300 W
	Imaging size	150 mm × 120 mm
	Pixel size	85 μm
	resolving power	58 Lp/ cm

The initial intensity of the incoming ray can be calculated using formula (3.7). When the ray passes through a gap of size, the intensity calculation formula is as follows:

$$P = P_0 e^{-\mu(x-\delta)} \quad (3.8)$$

where  $\delta$  represents the size of the gap between the conductor and ground wire. Based on the basic law of ray attenuation and the properties of exponential functions, the formula for the intensity of rays projected through various homogeneous materials can be obtained, as follows:

$$P = P_0 e^{-(\mu_1 x_1 + \mu_2 x_2 + \mu_3 x_3 + \dots + \mu_N x_N)} \quad (3.9)$$

Among them,  $x_N$  and  $\mu_N$  respectively represent the penetration thickness and attenuation coefficient of different objects [4].

During the detection process of the grounding wire, when the steel strand does not fully enter the steel pipe, the ray penetration thickness is the thickness of the aluminum pipe plus the thickness of the steel pipe, which is  $x_{\text{Aluminum tube}} + x_{\text{Steel pipe}}$ . When the steel strand fully enters the steel pipe, the ray penetration thickness is the thickness of the aluminum pipe plus the thickness of the steel pipe plus the thickness of the steel strand, which is  $x_{\text{Aluminum tube}} + x_{\text{Steel pipe}} + x_{\text{Steel strand}}$ . By comparing the ray penetration thickness when the steel strand does not fully enter the steel pipe and when it fully enters the steel pipe, and substituting the comparison results into formula (3.9), there is:

$$\begin{aligned} \frac{P_1}{P_2} &= \frac{P_0 e^{-(\mu_{\text{aluminium}} x_{\text{Aluminum tube}} + \mu_{\text{steel}} x_{\text{steel pipe}} + \mu_{\text{steel}} x_{\text{steel strand}})}}{P_0 e^{-(\mu_{\text{aluminium}} x_{\text{Aluminum tube}} + \mu_{\text{steel}} x_{\text{steel pipe}})}} \\ &= e^{-\mu_{\text{steel}} x_{\text{steel strand}}} \end{aligned} \quad (3.10)$$

Among them,  $P_1$  and  $P_2$  respectively represent the thickness of ray penetration when the steel strand does not fully enter the steel pipe and when it fully enters the steel pipe.

By comparing the radiation intensity at different times, the quality inspection of grounding wire crimping is achieved [12].

**4. Experimental analysis.** The model of the connecting tube grounding wire in this experiment is BLV450/750V, and the cross-section of the grounding wire is 5-35mm<sup>2</sup>. 1000 grounding wires are divided into 10 groups for every 100 wires. The steel core is automatically crimped at both ends of the grounding wire with a steel anchor size of 0-60mm, and the other end of the grounding wire with a steel anchor size of 70-100mm is automatically crimped. The manufacturer of the X-ray machine used in the experiment is Ruiao Testing Equipment (Dongguan) Co., Ltd., with the model RAYON-RE2100. The specific parameters are shown in Table 4.1 [10].

The accuracy of the crimping size of the grounding wire affects the crimping results. 10 sets of crimping sizes of the grounding wire were measured using the device described in this article, the traditional method 1

Table 4.2: Measurement results of ground ground (mm)

Group	Traditional method 1 device	Traditional method 2 device	The device	Actual measurement value
1	1.06	1.03	1.18	1.18
2	2.22	2.14	2.34	2.34
3	1.99	1.83	2.13	2.14
4	0.86	0.91	1.01	1.08
5	0.85	0.57	0.92	0.99
6	1.33	1.34	1.33	1.49
7	1.97	2.17	2.11	2.27
8	0.97	1.06	1.36	1.37
9	0.57	0.86	0.79	0.87
10	1.24	1.55	1.66	1.71
Average Discrepancy	0.24	0.20	0.06	/

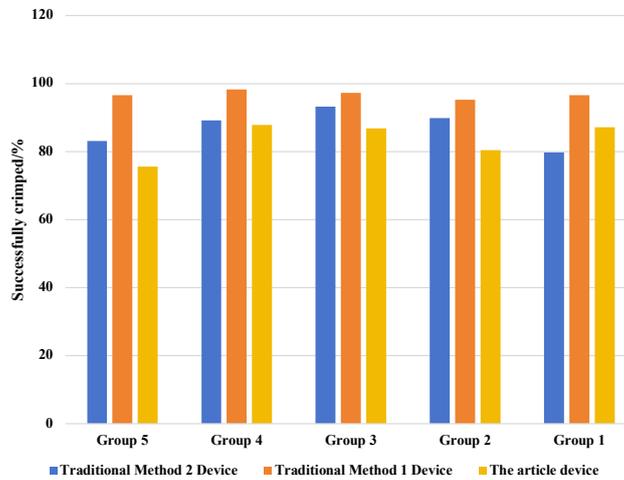


Fig. 4.1: Ground wire crimping results of three crimping devices

device, and the traditional method 2 device. The results are shown in Table 4.2, where traditional method 1 represents an electric drive hydraulic large-section grounding wire crimping device, and traditional method 2 represents a cable joint crimping device based on electromagnetic pulse forming technology.

According to Table 4.2, there is a certain deviation between the crimping dimensions of the grounding wires measured by the three devices and the actual measurement dimensions. Among them, the average deviation values of the 10 sets of measurement results for the traditional method 1 device and the traditional method 2 device are 0.24mm and 0.20mm, respectively. The difference between the crimping size of the grounding wire measured by the device and the actual measurement size is not significant, with an average measurement deviation of only 0.06mm, indicating that the device has high accuracy in the crimping measurement of the grounding wire [3].

Select 5 sets of grounding wires and use three crimping devices to crimp them, and verify the crimping results of the grounding wires. The results are shown in Figure 4.1.

Analyzing Figure 4.1, it can be seen that among the 5 sets of grounding wire crimping results, the success rate of traditional method 1 device grounding wire crimping is between 75% to 90%, and the success rate of traditional method 2 device grounding wire crimping is between 80% to 95%. The success rate of crimping the grounding wire of the device in this article is above 95%, and close to 100%. Compared with the traditional

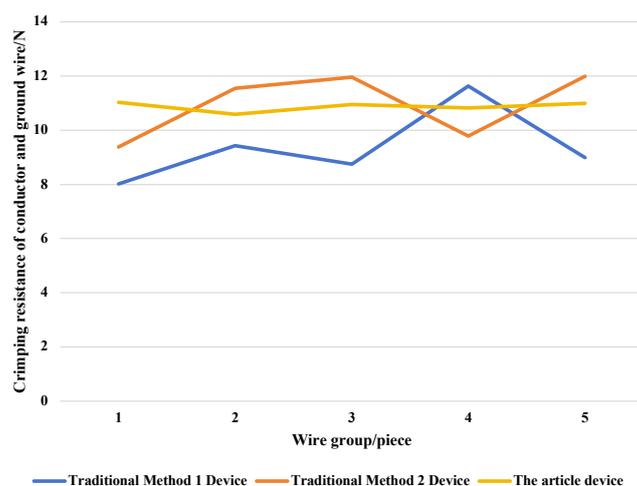


Fig. 4.2: Guide ground wire tensile test results

method 1 device and the traditional method 2 device, the device in this article has higher results in crimping the grounding wire [18, 9, 15].

During the automatic crimping process of the grounding wire, the crimping device must have high reliability and stability. Five sets of grounding wires are selected, and the crimping force is set to 11N. The information on the tension of the grounding wire during the crimping process is integrated, and the reliability and stability of the three crimping devices are analyzed. The results are shown in Figure 4.2.

Analyzing Figure 4.2, it can be seen that the tensile strength of the ground wire crimped by the traditional method 1 device and the traditional method 2 device differ significantly from the set tensile strength value. Among them, the maximum difference between the tensile strength of the ground wire crimped by the traditional method 1 device and the set tensile strength is 2N. Although there is some fluctuation in the tensile resistance value of the crimped grounding wire in the device, it is closest to the set tensile resistance value. Therefore, it can be concluded that the device has high reliability and stability.

Using sensitivity as a measure of the actual application performance of crimping devices, calculate the sensitivity of three crimping devices during the automatic crimping process of grounding wires, as shown in Figure 4.3.

Analyzing Figure 4.3, it can be seen that when the cross-sectional area of the grounding wire is different, the sensitivity of the device in this paper remains above 0.9, and the highest sensitivity reaches 0.98. The sensitivity of the traditional method 1 device and the traditional method 2 device varies from high to low, with the highest sensitivities of 0.92 and 0.88, respectively, which are significantly different from the sensitivity of the device in this paper. Therefore, it can be concluded that the sensitivity of the device in this paper is high and its practical performance is strong.

The inspection of the crimping quality of the grounding wire is an important step to ensure the quality of the grounding wire after crimping. The accuracy of the quality inspection of the grounding wire for three crimping devices is tested, and the results are shown in Figure 4.4.

Analyzing Figure 4.4, it can be seen that as the number of grounding wires increases, the detection accuracy curves of the traditional method 1 device and the traditional method 2 device show a downward trend. However, the detection accuracy curve of the device in this article is relatively smooth and the accuracy rate remains around 95%, indicating a high quality detection accuracy rate for grounding wires.

**5. Conclusion.** This article uses microcontroller and X-ray digital imaging technology to design an automatic crimping device for connecting tube conductors based on intelligent fully automatic technology. The microcontroller has powerful functions in data transmission, setting, clearing, testing, etc. It can not only handle any bit of special function registers but also perform logical operations on the bit. The device uses a

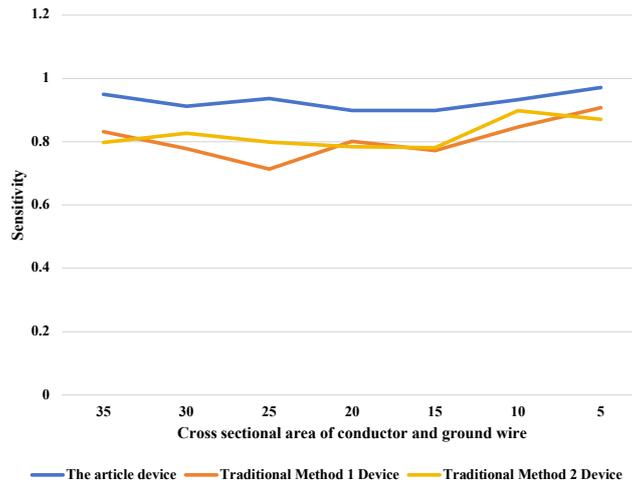


Fig. 4.3: Ground wire sensitivity of the three crimping devices

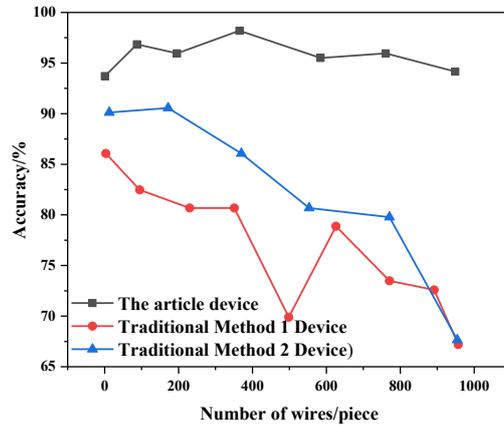


Fig. 4.4: Ground wire quality test results of the three crimping devices

microcontroller as a controller to achieve automatic crimping of the grounding wire. After experimental verification, it has been found that the difference between the measured crimping size of the grounding wire and the actual measurement size is small, with an average deviation of only 0.06mm; The success rate of crimping is above 95%, and the crimping results of conductor and ground wire are relatively high; The difference between the tensile strength and the set tensile strength of the crimped conductor wire is small, which has high reliability and stability; The highest sensitivity of the crimping device reaches 0.98, indicating a high sensitivity; The accuracy curve of ground wire quality detection is relatively smooth and the accuracy rate is maintained at around 95%.

REFERENCES

[1] A. AL TOUMA AND D. OUAHRANI, *The application of control algorithm for optimal performance of evaporatively-cooled façade system in hot dry and humid weathers*, Journal of Building Engineering, 44 (2021), p. 102642.  
 [2] Y. DENG, X. HE, G. ZHAO, K. ZHU, AND X. ZHANG, *Research on the application of self-optimizing control algorithm in tuned filters*, in International Conference on Computing, Control and Industrial Engineering, Springer, 2021, pp. 193–207.  
 [3] X. GUAN, L. HUANG, AND L. ZHANG, *Design of comprehensive service system for fresh food e-commerce under the background*

- of “rural revitalization”, *Journal of Physics: Conference Series*, 1757 (2021), p. 012141.
- [4] D. IZCI, *Design and application of an optimally tuned pid controller for dc motor speed regulation via a novel hybrid lévy flight distribution and nelder–mead algorithm*, *Transactions of the Institute of Measurement and Control*, 43 (2021), pp. 3195–3211.
  - [5] M. LAURINI, L. CONSOLINI, AND M. LOCATELLI, *A graph-based algorithm for optimal control of switched systems: An application to car parking*, *IEEE Transactions on Automatic Control*, 66 (2021), pp. 6049–6055.
  - [6] P. MAMOURIS, V. NASSIRI, G. MOLENBERGHS, M. VAN DEN AKKER, J. VAN DER MEER, AND B. VAES, *Fast and optimal algorithm for case-control matching using registry data: application on the antibiotics use of colorectal cancer patients*, *BMC medical research methodology*, 21 (2021), pp. 1–9.
  - [7] M. S. PACKIANATHER, T. ALEXOPOULOS, AND S. SQUIRE, *The application of the bees algorithm in a digital twin for optimising the wire electrical discharge machining (wedm) process parameters*, in *Intelligent Production and Manufacturing Optimisation—The Bees Algorithm Approach*, Springer, 2022, pp. 43–61.
  - [8] T. RIBEIRO, M. MASCARENHAS, J. AFONSO, H. CARDOSO, A. ANDRADE, S. LOPES, M. MASCARENHAS SARAIVA, J. FERREIRA, AND G. MACEDO, *P156 a multicentric study on the development and application of a deep learning algorithm for automatic detection of ulcers and erosions in the novel pillcam™ crohn’s capsule*, *Journal of Crohn’s and Colitis*, 16 (2022), pp. i232–i234.
  - [9] Y. SHANMUGAM, A. SUBBIYA, M. SUDHAKAR, AND D. H. VISALI, *Design of fir filter using adaptive lms algorithm for energy efficient application*, *Journal of Physics: Conference Series*, 1916 (2021), p. 012067.
  - [10] J. SHI, H. YAN, AND Y. WEI, *Application of the denoising technique based on improved fastica algorithm in cross-correlation time delay estimation*, *Journal of Physics: Conference Series*, 1754 (2021), p. 012196.
  - [11] E. SHINDIN, M. MASIN, A. ZADOROJNYI, AND G. WEISS, *On application of the simplex-type algorithm for scip to large-scale fluid processing networks*, arXiv preprint arXiv:2103.04405, (2021).
  - [12] H. S. TAVARES, L. BIFERALE, M. SBRAGAGLIA, AND A. A. MAILYBAEV, *Validation and application of the lattice boltzmann algorithm for a turbulent immiscible rayleigh–taylor system*, *Philosophical Transactions of the Royal Society A*, 379 (2021), p. 20200396.
  - [13] A. WADOOD, S. KHAN, B. M. KHAN, H. ALI, AND Z. REHMAN, *Application of marine predator algorithm in solving the problem of directional overcurrent relay in electrical power system*, *Engineering Proceedings*, 12 (2021), p. 9.
  - [14] J. WANG, G. CHEN, Z. ZHANG, J. SUO, AND H. WANG, *Wireless multiplexing control based on magnetic coupling resonance and its applications in robot*, *Journal of Mechanisms and Robotics*, 14 (2022), p. 011009.
  - [15] S. WANG, Y. SUN, C. YANG, AND Y. YU, *Advanced design and tests of a new electrical control seeding system with genetic algorithm fuzzy control strategy*, *Journal of Computational Methods in Sciences and Engineering*, 21 (2021), pp. 703–712.
  - [16] J. XU, L. XIAO, M. LIN, X. TAN, ET AL., *Application of fuzzy pid position control algorithm in motion control system design of palletizing robot*, *Security and Communication Networks*, 2022 (2022), pp. 1–11.
  - [17] S. YANZHE, L. DELIANG, J. SHAOFENG, L. YATING, AND W. HAO, *Review of the application of intelligent optimization algorithm for design of novel electric machines*, in *Proceedings of the 9th Frontier Academic Forum of Electrical Engineering: Volume II*, Springer, 2021, pp. 409–419.
  - [18] C. ZHANG, C. HU, S. XIE, AND S. CAO, *Research on the application of decision tree and random forest algorithm in the main transformer fault evaluation*, *Journal of Physics: Conference Series*, 1732 (2021), p. 012086.
  - [19] J. ZHANG, L. WU, J. HU, D. ZHAO, J. WAN, X. XU, ET AL., *Research and application of intelligent layout design algorithm for 3d pipeline of nuclear power plant*, *Mathematical Problems in Engineering*, 2022 (2022).
  - [20] X. ZHANG, G. YANG, Y. CUI, X. WEI, AND W. QIAO, *Application of computer algorithm in fault diagnosis system of rm equipment*, *Journal of Physics: Conference Series*, 2143 (2021), p. 012033.
  - [21] Z. ZHANG, *Application of ant colony optimization algorithm in the design of laser methane telemetry system*, in *International Conference on Multi-modal Information Analytics*, Springer, 2022, pp. 19–27.
  - [22] K. ZUO, Y. HONG, H. QI, Y. LI, AND B. LI, *Application of microwave transmission sensors for water cut metering under varying salinity conditions: Device, algorithm and uncertainty analysis*, *Sensors*, 22 (2022), p. 9746.

*Edited by:* B. Nagaraj M.E.

*Special issue on:* Deep Learning-Based Advanced Research Trends in Scalable Computing

*Received:* Sep 13, 2023

*Accepted:* Nov 4, 2023



## LINEAR ANTI-INTERFERENCE ALGORITHM FOR DIGITAL SIGNAL TRANSMISSION IN FIBER OPTIC COMMUNICATION NETWORKS BASED ON LINK ANALYSIS

JING WU\*, CHENG JIN† AND ZIWU WANG‡

**Abstract.** In order to achieve accurate transmission of protection signals in fiber optic communication networks, it is necessary to perform channel balancing configuration of fiber optic communication networks and adaptive forwarding control processing of relay protection signals, the author proposes an accurate transmission method for relay protection signals in fiber optic communication networks based on time-varying multipath fading suppression and adaptive beamforming. The system analyzes the sources of wireless long-distance pain signal interference signals, introduces anti-interference technologies such as two-dimensional joint processing (STAP), provides anti-interference algorithms and related gain analysis, and conducts signal processing gain simulation using MATLAB. Based on the analysis of comprehensive simulation results, at a given symbol length, the signal bandwidth increases, and the processing gain infinitely approaches the given theoretical limit value, rather than increasing nonlinearly. The reason is that the channel is affected by noise, and the channel estimation value and signal conjugate multiplication produce a noise quadratic term. At this point, the estimated value of the coherent region channel is reduced by the influence of noise, and the signal-to-noise ratio loss caused by the noise quadratic term is reduced, so the processing gain increases. During the process of infinite increase in signal bandwidth, the input signal-to-noise power ratio of the receiver tends to decrease towards an infinite value, limited by the size of the coherent region. The channel estimation value increases under the influence of noise, and the noise quadratic term is the main factor affecting the output noise power. When the symbol length is greater than the coherent time, the smaller the maximum Doppler frequency shift and the larger the coherent detection area, the greater the processing gain.

**Key words:** Wireless communication, Digital signal processing, Anti interference algorithm, Gain analysis, MATLAB simulation

**1. Introduction.** With the rapid development of wireless communication technology and fiber optic network communication technology, the use of fiber optic network communication to achieve large-scale data and signal transmission greatly facilitates people's production and work, and meets the needs of people for large bit sequence signal transmission and remote communication [6, 9, 3]. Fiber optic communication networks have the advantages of large transmission bandwidth and good real-time communication output. However, during the transmission of relay protection signals in fiber optic networks, multipath attenuation is prone to occur due to the influence of signal attenuation and transmission distance, moreover, the electromagnetic waves of relay protection signals are easily affected by environmental electric and magnetic fields, resulting in great ambiguity in the output of communication signals. Therefore, it is necessary to optimize the transmission design of relay protection signals in optical fiber communication networks, combining signal balance control and multipath attenuation suppression methods to improve the accurate transmission ability of relay protection signals in optical fiber communication networks, in traditional methods, there are mainly signal transmission control methods for relay protection signals in fiber optic communication networks, such as Baud interval equalization control, fractional interval equalization transmission control, beamforming method, and matched filter detection method. Frequency domain equalization method is used for inter symbol multipath suppression, and a fiber optic network communication channel equalization model is constructed to improve the accurate transmission ability of relay protection signals. Based on the above principles, Cao, Y. et al. proposed an energy internet data security transmission algorithm based on wireless opportunity network confidence to resist the impact of malicious behavior. The advantages of this algorithm in parameters such as success rate and data transmission delay were verified through simulation of real scenes [2]. Zhou, G. et al. proposed optical communication based on nonlinear Fourier transform (NFT) and digital coherent transceivers as a new theoretical framework for

---

\*Anhui Post and Telecommunication College, Anhui, Hefei, 230031, China (Corresponding Author).

†Anhui HaiXi Intelligent Co., LTD., Anhui, Hefei, 230000, China

‡Anhui Post and Telecommunication College, Anhui, Hefei, 230031, China

nonlinear fiber channel communication. For discrete eigenvalue transmission (or soliton transmission), people seek to encode as much information as possible in each degree of freedom and shorten the distance between adjacent pulses to improve the overall bit rate. However, this attempt can lead to nonlinear inter symbol interference (ISI) across multiple symbols and significantly reduce transmission performance [18].

On the basis of current research, the author proposes an accurate transmission method for relay protection signals in fiber optic communication networks based on time-varying multipath fading suppression and adaptive beamforming. The system analyzes the sources of wireless long-distance pain signal interference signals, introduces anti-interference technologies such as two-dimensional joint processing (STAP), provides anti-interference algorithms and related gain analysis, and conducts signal processing gain simulation using MATLAB [12]. Based on the analysis of comprehensive simulation results, at a given symbol length, the signal bandwidth increases, and the processing gain infinitely approaches the given theoretical limit value, rather than increasing nonlinearly. The reason is that the channel is affected by noise, and the channel estimation value and signal conjugate multiplication produce a noise quadratic term. At this point, the estimated value of the coherent region channel is reduced by the influence of noise, and the signal-to-noise ratio loss caused by the noise quadratic term is reduced, so the processing gain increases. During the process of infinite increase in signal bandwidth, the input signal-to-noise power ratio of the receiver tends to decrease towards an infinite value, limited by the size of the coherent region. The channel estimation value increases under the influence of noise, and the noise quadratic term is the main factor affecting the output noise power.

## 2. Methods.

### 2.1. Communication Channel Model and Relay Protection Signal Analysis.

(1) *Construction of communication channel model.* In order to achieve accurate transmission of relay protection signals in fiber optic network communication under strong interference, a fiber optic network communication channel model is first constructed. Fading channels are used for signal transmission link allocation and autocorrelation matched filtering design of the fiber optic network communication channel. It is assumed that the fiber optic network communication channel is a wide and stable channel with limited time domain bandwidth. The amplitude of the relay protection real signal  $x(t)$  received at the receiving end of the multipath fading channel between the received pulses is proportional to  $\alpha_k/N_k$  ( $\alpha_k$  is the amplitude attenuation of the  $k$ -th fading channel,  $N_k$  is the noise power spectral density), and the variance of channel transmission (variance) is defined as:

$$\sigma_x^2 = E [x^2(t)] \quad (2.1)$$

Among the  $n_R$  feedforward filters, the inter code interference intensity of the fiber optic network communication channel is  $s_i(t)$ ,  $i = 1, 2$ . Due to the varying path lengths of fiber optic network communication channels, the gain of the diversity equalizer branch satisfies the maximum likelihood sequence estimation (MLSE), and a feedback equalization model for fiber optic network communication channels is constructed, based on the maximum likelihood sequence estimation results, multipath propagation attenuation suppression of the communication channel is performed, and the calculation amount of each received symbol is proportional to  $M_s^N$ . The symbol transfer rate and measurement error of the fiber optic network communication channel are:

$$\begin{aligned} S_x &= E [x^3(t)] + \sqrt{s}bu [s(t - \tau_0)] \\ K_x &= E [x^4(t)] - 3E^2 [x^2(t)] b \end{aligned} \quad (2.2)$$

Among them,  $E[x^3(t)]$  is the attenuation feature of the relay protection signal symbols in each reception,  $b$  is the tap delay line, and the tap interval sampling method is used for sampling the transmission bit sequence, from this, the relay protection signal model outputted by the fiber optic network communication channel is obtained, and the instantaneous frequency estimation is performed based on the time-frequency distribution of the relay protection signal, thereby accurately simulating the pulse response of fiber optic network communication and wireless sensor network communication [7, 20]. The pulse response of the relay protection signal  $S(t)$  received by the fiber optic network communication system is:

$$S(t) = a_0 \sum_{i=1}^N a_i \delta(t - \tau_i) e^{jw_c t} \quad (2.3)$$

In the formula,  $N$  represents the number of communication channel paths in the fiber optic network,  $\tau_i$  and  $a_i$  respectively represent the time delay and frequency attenuation of the  $i$ -th fiber optic network communication channel, and  $\omega_c$  represents the output modulation carrier frequency of the relay protection signal in the fiber optic network communication, based on the time-varying multipath fading loss of the channel, the optimal baud interval sampling of the relay protection signal is carried out, and channel blind equalization technology is used to estimate the delay and amplitude of the communication channel. Establish a tap delay model for channel multipath component suppression, and set each node bi for fiber optic network communication, the pulse frame number of the channel's impact response is  $N_f$ . frames, and the output time delay in the phase offset direction is  $T_f$ . The pulse broadening of the relay protection signal in the fiber optic communication network is:

$$T_s = N_f T_f \quad (2.4)$$

In the transmission of relay protection signals in optical fiber communication networks, the multipath arrival time delay is  $T_d$ , and the relevant beam modulation method is used, divide the relay protection signal in the fiber optic communication channel of the fiber optic network into  $N_c$  chips, with a tap time interval of  $R_b < 1/\Delta$ , and obtain the time extension of the relay protection signal output that meets the following requirements:

$$T_c = \text{ent}(T_f/N_c) \quad (2.5)$$

It can be seen that there is a spectral zero in the frequency response of the channel at the communication output end. Through channel equalization allocation, a phase shift is generated in the received signal, improving the accurate transmission performance of relay protection signals in fiber optic network communication, and thus improving communication quality [14, 5].

(2) *Time-varying multipath attenuation suppression.* Based on the above construction of the fiber optic network communication channel model, the optimal baud interval sampling of the relay protection signal is carried out based on the time-varying multipath fading loss of the channel. The IEEE802.15.SG3a communication protocol is used to construct the shortest routing algorithm for the relay protection signal transmission channel of fiber optic communication. The node distribution model of the fiber optic communication data link layer is constructed, and the output sampled relay protection signal model is obtained as follows:

$$s(t) = \sum_i b_j \sum_{j=0}^{N_f-1} p(t - iT_s - jT_f - c_j T_c) \quad (2.6)$$

Assuming that the impulse response of the relay protection signal in the fiber optic network communication channel is  $h(n)$ , the nonlinear equilibrium parameter of the relay protection signal transmission channel is  $n(n)$ , and the time-varying multipath fading loss in the frequency domain is  $y(n)$ , the symbol of each feedforward branch is  $\tilde{x}(n)$ . Under a limited symbol rate, the optimal baud interval sampling of the relay protection signal is performed based on the time-varying multipath fading loss of the channel. The spectrum characteristics of the relay protection signal in the output fiber optic network are:

$$\text{Computation}(n_j) = [(E_{\text{elec}} + E_{\text{DF}}) \delta + E_{\text{elec}} + \varepsilon_{fs} d_j^2] l \quad (2.7)$$

Adjust the tap coefficient of the fiber optic network communication system, use blind equalization method for multipath suppression, and improve the output fidelity of the signal [4, 11].

**2.2. Analysis of Interference Signals and Anti interference Techniques.** Interference signals often come from potential interference and adversarial interference sources such as natural and human interference sources. The sources of natural environmental interference are the astronomical noise generated by charges in the atmosphere, the cosmic noise in outer space of the Earth, and the background noise of the natural environment [13]. Human interference sources mainly come from various electromagnetic interferences that humans believe to be generated, including signal noise interference generated by active radiation devices such as broadcasting, television, radar signals, and mobile communication; At the same time, there are various

high-voltage transmission lines and engineering equipment that emit electromagnetic interference. Interference with signal reception is extremely common. In response to communication conditions in harsh environments, spread spectrum technology serves as an effective anti-interference technology in wireless signal transmission communication. Spread spectrum can reduce interception probability and increase communication distance. The pseudo random sequence signal used for spread spectrum has white noise statistical characteristics, and the spread spectrum signal can effectively combat multipath fading characteristics. With the development of technology, two-dimensional spread spectrum technology has been adopted, that is, synchronous spread spectrum in the time and frequency domains, fully utilizing the signal's correlation gain in the time and frequency domains to improve its anti-interference ability. Due to the surge in wireless communication traffic and the increasing complexity of long-distance communication transmission environments, as well as the need for high frequency spectrum utilization in multi-system communication requirements. We have started to adopt multi antenna multi interface technology (MIMO), which has array gain and coding characteristics, which can effectively suppress transmission channel interference, significantly increase data transmission rate, and expand communication capacity [16, 8].

In recent years, technicians from various countries have adopted technologies such as Orthogonal Frequency Division Multiplexing (OFDM), Orthogonal Frequency Division Multiplexing Multiple Access (OFDMA), and Single Carrier Frequency Division Multiplexing (SC-FDMA). It effectively solves the problem of spectrum resources. With the improvement of broadband wireless data services such as wireless local area networks and wide area networks, the Wireless Ethernet Compatibility Alliance (WECA) has proposed WiFi technology, Bluetooth technology, and Zigbee technology applied to multi-sensor networks such as the Internet of Things, creating a more complex wireless communication environment today.

### 2.3. Space Time Adaptive Anti-interference Technology.

(1) *Two dimensional joint processing (STAP) anti-interference technology*. At present, space-time adaptive anti-interference technology is the best technology to effectively ensure anti-interference reception of satellite signals in complex electromagnetic environments. This technology has extremely high dynamic characteristics and can quickly change the spatial response based on changes in the electromagnetic environment, ensuring stable reception of satellite navigation signals [17]. The main principle is to use the antenna array and time-domain delay to utilize the spatial and energy characteristics of the interference signal, and achieve spatial interference to zero, thereby achieving interference signal cancellation and suppression. Space time adaptive anti-interference technology has the characteristics of small size, modular design, and strong expansion compatibility. Space time adaptive anti-interference mainly includes time domain adaptive anti-interference and spatial domain adaptive anti-interference. Time domain adaptation is based on FIR filters, which achieve frequency domain filtering through time domain delay to filter out non operating frequency signals, and achieve maximum suppression of interference signals through spatial and temporal synchronization cooperation. The spatial adaptive anti-interference mainly relies on the spatial characteristics of the interference signal to achieve zero adjustment function for the interference direction in the spatial directional map. Through the combination of time-domain filtering, frequency-domain filtering, and spatial filtering, a spatiotemporal two-dimensional joint processing (STAP) anti-interference technology is formed. STAP technology extends one-dimensional time-domain, frequency-domain, and spatial filtering to the two-dimensional domains of time and space, forming a spatiotemporal two-dimensional processing structure. The combination of digital signal processing and array technology completes the adaptive processing function of the system. The spatiotemporal two-dimensional joint adaptive processing structure is derived using statistical likelihood ratio detection theory under the model of overlaying deterministic signals in the background of Gaussian noise, it uses a multi tap FIR filter in each element structure to achieve anti-interference processing. In the channels of each array element, various levels of delay constitute FIR filters to filter out interference in the time domain; Different time delay nodes of array elements form adaptive filtering in the spatial domain, which distinguishes spatial interference sources and forms spatial nulls in the spatial domain to suppress interference. Space time processing can also achieve the function of eliminating interference in the two-dimensional space time domain. Figure 2.1 shows a spatiotemporal frequency domain adaptive anti-interference loop.

(2) *STAP anti-interference algorithm*. The spatiotemporal two-dimensional joint processing technology lies in how to implement the spatial weight vector solving method. The time and space weight vectors are con-

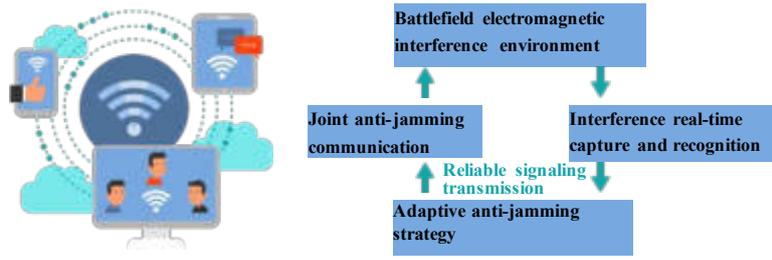


Fig. 2.1: Space time frequency domain adaptive anti-interference loop

strained by various criteria such as the minimum mean square error criterion, the maximum signal-to-noise ratio criterion, the linearly constrained minimum variance criterion, and the maximum likelihood criterion. In practical engineering applications, appropriate constraint criteria are selected based on different environments [15, 19]. The Linear Constrained Minimum Variance (LCMV) criterion is widely used. It utilizes the characteristics of navigation signal power being much lower than noise and interference power to weaken interference energy and avoid the impact of navigation signals. When it is difficult to separate signals from interference signals, the optimal processor of LCMV ensures that the signal is lossless by constraining both the spatial and temporal domains under the condition of spatiotemporal joint processing, adjusting the weights to minimize the variance of the output signal (power). The mathematical algorithm formula for the LCMV criterion is:

$$\min_W \{W^H R_X Y\} \text{ s.t. } W^H C_X = g \tag{2.8}$$

In the above equation,  $R_x$  represents the covariance matrix of the received signal,  $R_X = E(XX^H)$ .  $C_x$  represents the constraint matrix;  $G$  represents the constraint response vector. The spatial weight vector is solved using the Lagrangian equation:

$$W_{\text{opt}} = R_X^{-1} C_X (C_X^H R_X^{-1} C_X)^{-1} g^H \tag{2.9}$$

The linear constraint minimum variance criterion is limited by  $C_x$ , and its application conditions depend on the changes in the constraint matrix [1].

*Gain analysis.* The time-frequency two-dimensional spread spectrum in multipath fading and fully scattering channels is a BPSK modulated signal, for the joint coherent and incoherent detection algorithm, the spread spectrum signal processing gain adopts the ratio of the output signal-to-noise ratio of the signal receiver to the input signal-to-noise ratio.

$$G = \frac{p_{so}/p_{no}}{p_{si}/p_{ni}} \tag{2.10}$$

In the above equation,  $P_{so}$  represents the output signal power and  $P_{no}$  represents the output noise power. Considering the influence of channel fading factor, which is a complex Gaussian random variable, the output signal-to-noise ratio of the signal after joint coherent and incoherent detection is:

$$\frac{P_{so}}{p_{no}} = \frac{1}{4} \frac{B^2 T^2 (1 + N_B N_T) (P_{si}/P_{ni})^2}{B T P_{si}/P_{ni} + 1} \tag{2.11}$$

Based on the above formula calculation and analysis, the more coherent regions participate in non coherent detection, the higher the output signal-to-noise ratio. Considering the time-frequency twodimensional spread spectrum signals in multipath fading channels, for joint coherent and incoherent detection algorithms, the independence of coherent regions and the corresponding increase in reception diversity are beneficial for increasing

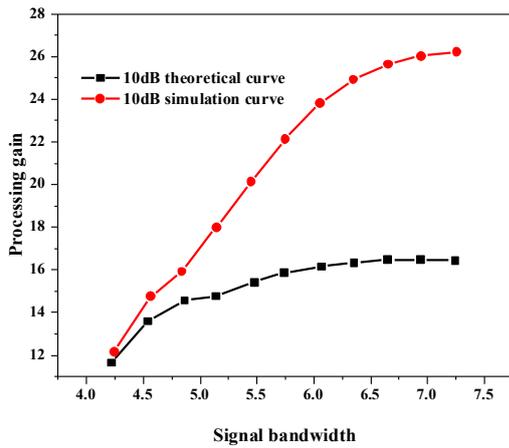


Fig. 3.1: Comparison of 10dB theory and simulation

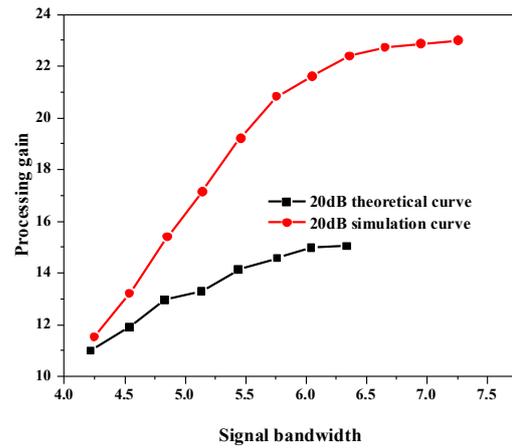


Fig. 3.2: Comparison of 20dB theory and simulation

the signal-to-noise ratio of the output signal in the detection algorithm. Simplifying the above equation, the spread spectrum signal processing gain is:

$$G = \frac{1}{4} \frac{B^2 T^2 (1 + N_B N_T) P_{si} / P_{ni}}{B T P_{si} / P_{ni} + 1} \tag{2.12}$$

The above analysis shows that when the input signal-to-noise ratio is large enough, the processing gain is directly proportional to the size and number of coherent regions. When the input signal-to-noise ratio is small enough, the output signal-to-noise decreases rapidly, and the processing gain is directly proportional to the size and number of coherent regions, as well as the input signal-to-noise ratio.

**3. Results and Analysis.** Using MATLAB simulation software, simulate the gain of time-frequency two-dimensional spread spectrum signal processing. The simulation results are shown in Figures 3.1, 3.2, 3.3, and 3.4. During the simulation process, the channel response in the coherent region is not completely correlated, with a time domain correlation coefficient not greater than 0.5 and a frequency correlation coefficient not greater than 0.9. The simulated output signal-to-noise and processing gain values are lower than the theoretical values [10].

Based on the analysis of comprehensive simulation results, at a given symbol length, the signal bandwidth increases, and the processing gain infinitely approaches the given theoretical limit value, rather than increasing nonlinearly. The reason is that the channel is affected by noise, and the channel estimation value and signal conjugate multiplication produce a noise quadratic term. At this point, the estimated value of the coherent region channel is reduced by the influence of noise, and the signal-to-noise ratio loss caused by the noise quadratic term is reduced, so the processing gain increases. When  $\frac{E_b}{N_0}$  is constant, the signal bandwidth increases infinitely, and the input signal-to-noise power ratio of the receiver decreases towards an infinite value. Limited by the size of the coherent region, the channel estimation value increases under the influence of noise, and the noise quadratic term is the main factor affecting the output noise power. When the symbol length is greater than the coherent time, the smaller the maximum Doppler frequency shift and the larger the coherent detection area, the greater the processing gain.

**4. Conclusion.** Optimize the transmission design of relay protection signals in fiber optic communication networks, combining signal balance control and multipath attenuation suppression methods, to improve the accurate transmission ability of relay protection signals in fiber optic communication networks, the author proposes an accurate transmission method for relay protection signals in fiber optic communication networks based on time-varying multipath fading suppression and adaptive beamforming. The wireless communication system has limited transmission power, long-distance signal propagation, harsh and complex channel environments,

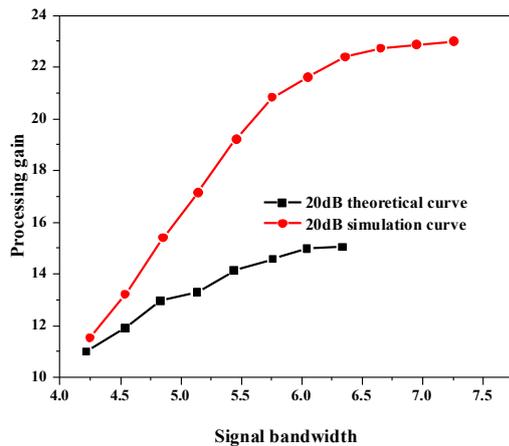


Fig. 3.3: Comparison of 10Hz theory and simulation

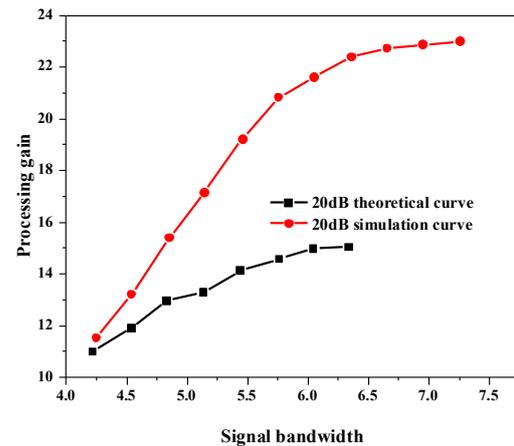


Fig. 3.4: Comparison of 00Hz theory and simulation

extremely low signal-to-noise ratio, increasingly complex wireless communication network structure, and severe interference from physical and system environments on the receiving end signal. The author proposed a two-dimensional joint processing (STAP) anti-interference technology to analyze and simulate the gain of the joint coherent and incoherent detection algorithm for time-frequency two-dimensional spread spectrum modulated signals in multipath fading and fully scattered channels using MATLAB. The results showed that the simulated output signal-to-noise and processing gain values were lower than the theoretical values.

#### REFERENCES

- [1] M. K. ALSHAWAQFEH, O. S. BADARNEH, AND F. S. ALMEHMADI, *Performance analysis of digital transmission in interference-limited networks*, Telecommunication Systems, 79 (2022), pp. 233–247.
- [2] Y. CAO, X. LI, J. LIU, C. LI, J. YAN, AND J. ZHAO, *Research on secure data transmission algorithm based on confidence of wireless opportunistic networks for energy internet*, 2087 (2021), p. 012094.
- [3] X. CHEN, J. XU, X. ZHANG, AND W. ZHANG, *Research on the filter algorithm for integrated autonomous navigation based on angle and velocity measurement in deep space*, Scientia Sinica Physica, Mechanica & Astronomica, 52 (2022), p. 214509.
- [4] Y. DAI, J. LIU, M. SHENG, N. CHENG, AND X. SHEN, *Joint optimization of bs clustering and power control for noma-enabled comp transmission in dense cellular networks*, IEEE Transactions on Vehicular Technology, 70 (2021), pp. 1924–1937.
- [5] S. GONG, J. ZHANG, S. MI, F. FAN, AND B. YAN, *Near-ml detection based on multilevel sinr thresholding for mode division multiplexing transmission with mode-dependent loss*, IEEE Access, 9 (2021), pp. 20136–20142.
- [6] Y. GUAN, Z. HU, C. CHEN, AND X. ZHU, *An anti-noise transmission algorithm for 5g mobile data based on constellation selection and channel joint mapping*, Alexandria Engineering Journal, 60 (2021), pp. 3153–3160.
- [7] J. LI, Y. LIANG, H. WANG, AND H. XIN, *Research on efficiency improvement strategy of optical fiber signal transmission based on fourier transform*, 1744 (2021), p. 022048.
- [8] Z. LIU, C. LIANG, Y. YUAN, K. Y. CHAN, AND X. GUAN, *Resource allocation based on user pairing and subcarrier matching for downlink non-orthogonal multiple access networks*, IEEE/CAA Journal of Automatica Sinica, 8 (2021), pp. 679–689.
- [9] J. LV, C. QU, S. DU, X. ZHAO, P. YIN, N. ZHAO, AND S. QU, *Research on obstacle avoidance algorithm for unmanned ground vehicle based on multi-sensor information fusion*, Mathematical Biosciences and Engineering, 18 (2021), pp. 1022–1039.
- [10] Z. SHEN, K. XU, AND X. XIA, *Beam-domain anti-jamming transmission for downlink massive mimo systems: A stackelberg game perspective*, IEEE Transactions on Information Forensics and Security, 16 (2021), pp. 2727–2742.
- [11] F. SHI, Y. FAN, X. WANG, Q. TAN, AND Y. GAO, *Rf self-interference cancellation for radio-over-fiber link based on dual phase modulation in sagnac loop*, IEEE Photonics Journal, 13 (2021), pp. 1–15.
- [12] X. WANG, X. DENG, H. SHEN, G. ZHANG, AND S. ZHANG, *Speech intelligibility enhancement algorithm based on multi-resolution power-normalized cepstral coefficients (mrpncc) for digital hearing aids.*, CMES-Computer Modeling in Engineering & Sciences, 126 (2021), pp. 693–710.
- [13] Y. XU, J. CHENG, G. WANG, AND V. C. LEUNG, *Adaptive coordinated direct and relay transmission for noma networks: A joint downlink-uplink scheme*, IEEE Transactions on Wireless Communications, 20 (2021), pp. 4328–4346.
- [14] N. ZHANG, *Retracted: Intelligent control optimization of remote transmission quality of communication signals based on 5g network*, 2066 (2021), p. 012031.

- [15] N. ZHANG, B. CHEN, Y. YANG, H. SUN, M. CHEN, T. CHEN, X. CHEN, M. XIAO, AND Y. LUO, *Optimization technology of optical fiber communication network based on service classification*, 1746 (2021), p. 012085.
- [16] X. ZHANG AND J. LI, *Power control for cognitive users of perception layer in complex industrial cps based on dqn*, IEEE Access, 9 (2021), pp. 25371–25382.
- [17] F. ZHAO, Y. JIANG, Y. WANG, AND J. TANG, *Tcpw-f: A congestion control algorithm based on noise-aware in wireless networks*, 1995 (2021), p. 012039.
- [18] G. ZHOU, L. SUN, C. LU, AND A. P. T. LAU, *Multi-symbol digital signal processing techniques for discrete eigenvalue transmissions based on nonlinear fourier transform*, Journal of Lightwave Technology, 39 (2021), pp. 5459–5467.
- [19] L. ZHOU AND J. YU, *Resource allocation for high-speed train communication based on deep reinforcement learning*, 1827 (2021), p. 012184.
- [20] Y. ZHOU AND H. SU, *Research on mode-coupled multi-mode fiber mode-division in multiplexing transmission system*, 2246 (2022), p. 012034.

*Edited by:* B. Nagaraj M.E.

*Special issue on:* Deep Learning-Based Advanced Research Trends in Scalable Computing

*Received:* Sep 13, 2023

*Accepted:* Nov 6, 2023



## NETWORK TRAFFIC MONITORING AND REAL-TIME RISK WARNING BASED ON STATIC BASELINE ALGORITHM

ZHAOLI WU\* AND JUNWEI LIU†

**Abstract.** With the rapid growth of network traffic, in order to monitor network traffic, the author proposes a baseline based traffic inspection method. The main objective is to develop a global system for identifying malicious traffic, rather than a precise method for detecting the types of worms produced by malicious traffic. Although traffic is caused by the causes, network administrators can use this international search technique to detect malicious traffic data. The system based approach mainly includes designing time based on the traditional traffic model, detecting various equipments and network traffic process, and configuring the traffic flow according to each time frame. This method uses Cisco's NetFlow Collector, a NetFlow Collector (NFC), to collect raw NetFlow data transmitted by the device through UDP every 5 minutes. the Then, three-dimensional data such as communication port, communication time, and traffic flow (bytes or packets) is used to filter, remove the different values, calculate the base values, and compare the real-time results with the base values to check the traffic defects in the current network. If there are differences between the monitoring data and the system configuration at the same time, the system will issue an abnormal warning, and as time accumulates, the alarm level will gradually escalate.

**Key words:** Static baseline algorithm, Network traffic monitoring, Real time performance, Risk warning

**1. Introduction.** With the rapid development of broadband internet in China, the network scale of major operators in the country is constantly expanding, the network structure is becoming increasingly complex, network services are becoming increasingly rich, network traffic is growing rapidly, and the network environment is becoming increasingly complex [16]. Operators need to use reliable and effective network traffic monitoring systems to conduct timely and accurate traffic and flow analysis of their networks and various services carried by them, in order to tap into the potential of network resources, control network interconnection costs, and provide a basic basis for network planning, optimization and adjustment, and business development. Mainly manifested in the following aspects: By analyzing the network outlet flow and direction, it is possible to gain a detailed understanding of the access of internal users to other external networks, thereby effectively selecting the interconnection method and location with other networks, and saving interconnection link costs. Master the user's access to other operators: By monitoring the traffic of Wulian with other networks, analyze the business characteristics and main traffic directions of internal users accessing other external networks, accurately grasp the interests of internal users on the external network, and find the most applied hot information content. Based on the analysis results, the corresponding network content is constructed, and the hot information content that users are interested in is placed in the internal network to reduce the pressure on the interconnection link. Evaluate the cost and value of branch networks: By monitoring the incoming and outgoing traffic of each branch network, analyzing the size, direction, and content composition of the traffic, we can understand the bandwidth usage of each branch network, reflect the network cost it occupies, and also understand its business development and make a value evaluation [10, 9]. IP based billing applications and Service Level Agreement (SLA) verification services: By monitoring and analyzing the traffic on major customer access circuits, parameters such as business type, service level, communication time and duration, and communication data volume can be calculated, providing data basis for IP based billing applications and SLA verification services. By monitoring specific traffic in the network for a long time, it helps network administrators understand

---

\*1. Jiangsu Vocational Institute of Architectural Technology School of Information and Electronics Engineering, Yangzhou, Jiangsu, 225006, China; 2. Jiangsu Collaborative Innovation Center for Building Energy Saving and Construction Technology, 221000, China; 3. School of computer science and technology, China University of mining and Technology, Xuzhou, Jiangsu, 221000, China. Corresponding author, [ZhaoLiWu5@126.com](mailto:ZhaoLiWu5@126.com)

†School of Internet of Things Technology, Wuxi Vocational College of Science and Technology, Wuxi, Jiangsu, 214028, China. [JunweiLiu6@163.com](mailto:JunweiLiu6@163.com)

the network's traffic model. The benchmark data formed can be used by network administrators to correctly analyze the network usage status, and timely issue abnormal alarms. Preventive measures can be implemented before fault events occur or expand, thereby improving the overall quality and efficiency of the network [3]. The implementation of communication network vulnerability analysis focuses on the prevention of denial of service (DDoS) attacks and the spread of major diseases. Real-time analysis of network traffic quality helps to identify network traffic quality in time and identify the specific characteristics of network traffic quality. By comparing with the traditional principle of network communication, administrators can quickly determine whether the abnormal communication is the protection of network security, determine the type of security precautions, measure the potential danger and the potential to affect various kinds of attacks, develop and implement anti-accident measures in an emergency system. and implement anti-accident measures. By analyzing the traffic flow, the data base can be provided for network performance evaluation such as multi-outlet load balance, critical link bandwidth location, route selection, QoS location, etc. the network performance evaluation method can be used to evaluate the network performance of the network system. With the development of computer and communication technology, the scale and complexity of electric power enterprises are increasing, and more and more industrial applications are also becoming more and more complex. As a result, the possibility of various types of traffic jams is also increasing, and various traffic jams are also following. The generation of traffic jam not only affects the performance of network and reduces efficiency, but also can increase information confidentiality of enterprises, affect their development [17]. Therefore, differential detection has become an important challenge that power companies are facing. The main flow recording technologies mainly include monitoring protocol technology based on mirror (valve online) flow, distributed monitoring technology (based on hardware analysis technology), NetFlow based on monitoring technology, and SNMP based on monitoring technology. Compared to NetFlow monitoring technology, other technologies also have their own disadvantages. Automobile telescope can only be used for single link and is not suitable for wide area network inspection. The hardware probe technology is limited by upper limit of intersection speed. SNMP technology mainly collects some status information of equipment and related information. Because of the monotonicity of the data and some errors, data analysis is only possible on the data of network layer 2 and 3 and the status of the future equipment. NetFlow monitoring technology is involved in a unified process that does not rely on certain links. At the same time, it has high efficiency in collecting data, a wide range of network applications, lower costs compared to others, and high cost-effectiveness. In response to this research issue, Gong, Q. et al. a passive, nonautonomous intrusion detection system for 100G research networks is proposed. Lightning safety technology uses multi-core and GPU to achieve high quality packet and pipeline detection. It extracts the physical and physical conditions of the real time data in the network and gives them to the network anomaly detection system [4]. Bollmann, C. A. et al. this paper describes a new application of robust estimation in fault diagnosis of computer network traffic volume. The proposed algorithm is based on the position and distribution of the model, and is derived from the unknown zero order numbers. This test is not parametric and suitable for large-scale data analysis of external traffic network. Using two different global denial of service attacks including the actual ability of backbone network traffic to verify the effectiveness of these measures. Because of the heavy tails in the network traffic system (a special kind of real network traffic system), the simulation results are better than the traditional test means such as interpolation and interpolation. Monte Carlo analysis was used to evaluate the efficacy and apparent improvement rate from 7% to 11% in false negative. The test request is also shown to be associated with an average (a similar test case) [2].

On the basis of current research, the monitoring method proposed by the author, although lacking in accuracy and unable to fully monitor abnormal traffic, has been greatly simplified due to its highly singular mechanism. It is very useful for a complex and large network of a large enterprise, with improved efficiency and better real-time performance.

## 2. Methods.

### 2.1. Static baseline algorithm.

(1) *Definition and Principle.* The 'static baseline algorithm' refers to setting the same upper limit (upper baseline) or lower limit (lower baseline) that does not change over time within a 24-hour cycle range for a certain indicator, and dividing the normal range and abnormal area of the indicator value [20]. The algorithm principle is as show in Figure 2.1.

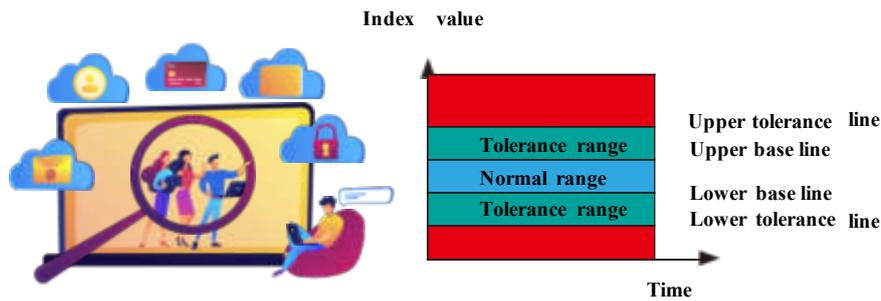


Fig. 2.1: Static baseline algorithm

For certain indicators, there may only be upper tolerance line or lower tolerance line, such as CPU load; Other indicators may require simultaneous attention, with both the upper and lower tolerance lines reflecting abnormal situations. The upper and lower tolerance line settings for distinguishing between normal and abnormal severity levels come from network operation and maintenance experience, management requirements, or device capacity limitations, and the accuracy of the settings determines whether the algorithm can work. Excessive tightness (such as a small upper limit value or a large lower limit value) may lead to false positives [6, 15]; Excessive looseness (such as high upper limit values and low lower limit values) may lead to false positives, increase the workload of monitoring personnel, affect the enthusiasm of maintenance personnel, and ultimately reduce the effectiveness of this technology. At the same time, as the number of indicators, network elements, and business systems included in the active monitoring scope increases, the number of upper and lower tolerance lines that need to be set will also increase sharply, indicating a practical need to improve work efficiency. Therefore, on the premise of pursuing accuracy, finding the automatic threshold setting mode becomes an important factor for the successful application of this algorithm, which needs to be constantly adjusted and optimized in the operation and maintenance.

(2) *Method for determining the baseline.* Traditionally, maintenance personnel determine thresholds for various indicators based on experience and set them manually in network management systems and other means, abbreviated as the “manual setting” method [19]. This method can adapt to work needs when there are few indicators and small fluctuations. However, with the inclusion of active monitoring indicators and a large number of equipment, this method is difficult to adapt to operational needs and has low work efficiency; For indicators with large fluctuations, it is easy to have arbitrary settings and strong subjectivity, which is not conducive to system maintenance. Of course, manual setting can absorb the experience of operation and maintenance personnel, and has a certain degree of flexibility. The indicators related to management requirements and equipment capabilities applicable to static baseline algorithms can be manually set to monitor whether they are below the assessment value or exceed the corresponding proportion of processing capacity; Other indicators applicable to this algorithm, due to their certain fluctuation range and compliance with certain statistical laws, can be automatically learned by the system to generate corresponding thresholds based on the principle of dynamic baseline algorithm, and automatically set. This method is referred to as the “automatic setting method”. Therefore, both methods are applicable to different types of indicators, with the latter having a wider scope of application.

**Manual setting method:** Maintenance personnel refer to the historical performance, management requirements, or equipment capabilities of a certain indicator, and combine their own work experience to set the baseline: manually write it into the system configuration file or fill it in the configuration window to set the baseline [5].

**Automatic setting method:** In summary, the “automatic setting method” adopts a principle similar to the dynamic baseline algorithm, based on historical statistical data, regardless of time period differences, calculates a threshold that does not change with 24 hours of time, and automatically sets it as the baseline of normal data. The algorithm description is as follows: historical data values; A total of  $N$  pieces of data within all

Table 2.1: NetFlow field description

byte	content	describe
0~3	srcaddr	IP address of the source
4~7	dstaddr	IP address of the destination
8~11	ne: xthop	IP address of the next network segment router
12~15	input and output	SNMP index for input and output interfaces
16~19	dPkts	Packets in this information flow
20~23	dOctets	The total number of Layer3 bytes in a packet of information flow
24~27	First	SysUptime at the beginning of information flow
28~31	Last	SysUp time when the last packet of the information flow is received
32~35	srcport and dstport	TCP/UDP source and destination port numbers
36~39	pad1 ,prot and tos	Unused (i.e. content 0) bytes, IP protocol (e.g. 6=TCP, 17=UDP), and IP service type
40~43	Flags, pad2, pad3	The cumulative OR of TCP flags, pad2 and pad3 are unused (i.e. bytes with content 0)
44~48	reserved	Unused (i.e. bytes with content 0)

collection granularity for M consecutive days: If the collection granularity is 15 minutes, then the number of data  $N=MX24X$  (60/15); If the collection granularity is 30 minutes, then the number of data  $N=MX24X$  (60/30). N data are denoted as  $x1\sim xN$ . Exclude abnormal data; Based on the maintenance records during this statistical period, eliminate data during holidays, malfunctions, and errors. It is also possible to arrange a set of data in a fixed proportion. When the system is implemented, the probability of normal data can be manually adjusted and then calculated. Determine the range of baseline values; Priority should be given to the probability distribution algorithm in the dynamic baseline algorithm section, followed by the sorting method.

**2.2. NetFlow introduction.** Cisco's NetFlow is a switching technology introduced under the IOS system [14]. It utilizes seven attributes with segment identification, including source IP address, target IP address, TOS byte, and layer 3 protocol type. At the same time, while being supported by Cisco routers, it quickly and accurately distinguishes different types of Flow, and then tracks, measures, and analyzes them to obtain information such as time, type, and size. NetFlow services can provide some other efficient and high-quality services, such as data statistics under the optimal exchange path cooperating with routers, efficient data statistics under the maximum limit of information interaction between routers and switches, and multi type data statistics such as users, protocols, ports, and service types. In addition, NetFlow can also be deployed anywhere in the network as a pathable device, and to some extent, security services can be implemented through packet filtering.

NetFlow has a relatively small impact on routers due to its use of special switching technology [18]. The query process only targets the first group. After the Flow is identified and distinguished, the subsequent groups default to being a part of it. This is oriented towards connection based processing, avoiding the operation of access lists, and thus achieving data collection under small impact. By further analyzing the field descriptions of NetFlow, its superiority can be better reflected. As shown in Table 2.1, it can be found that the flow record contains information with data flow identification such as the source/destination IP address, transport layer source, destination port number, etc., which includes tcpflags for data security related data. Address related data such as source address, destination address, source autonomy number, and destination autonomy number, etc. NetFlow provides many details about the data, it can analyze various information of data in detail.

In summary, NetFlow technology, due to its unique formal architecture, does not repeatedly search for each data packet, but analyzes the data flow, improving the previous inefficient traffic monitoring and high impact on devices, making it more suitable for large networks.

**2.3. Baseline analysis.** Basis analysis is a multi- period averaging algorithm which deals with the sequence of time, calculates the average value of the variables in each phase, and links their results to make a practical basis. It represents the normal value of the system, and when used for vehicle maintenance, it intro-

Table 2.2: Description of abnormal types of network abuse

Abnormal	Definition	Illustrate
ALPHA	Unusual high byte rate transmission from point to point	Bandwidth measurement experiment
DOS,DDOS	(Distributed) Denial of Service Attack	A large number of data packets are sent to a specific port of a separate destination IP (such as port 0)
Flashing crowding	Unusual high volume of requests for resources or services	Large number of web requests for a single IP (port 80)
scanning	Scan a host for vulnerable ports (port scanning) or scan the network for an attack target (network scanning)	Network Scan for Port 139 (NetBIOS)
worm	Autobroadcast code can be spread across the network through security vulnerabilities	Attack via Port1433 (MSSQL Snake worm)
Point to multipoint	Content distribution from one server to many users	A single server broadcasts to a large number of destination address sets on port 119
Transmission loss	Traffic reduction event caused by traffic exchange between OD flow pairs	Shutdown of nodes on the scheduled network backbone and measurement failure of nodes
Entrance transfer	Consumers transfer traffic from one entrance to another	Online consumers transfer their business traffic from one network to another

duces a traditional vehicle maintenance system [8]. The author adopts a baseline analysis method to monitor: certain specific ports (blacklisted) of well-known ports (port numbers below 1024), namely some published ports related to abnormal traffic, such as some worm viruses (135 ports used by shockwave viruses); All other non well-known ports (port numbers above 1024), excluding ports for certain specific applications (such as some internal communication software of enterprises) (whitelist method). By conducting baseline analysis on the historical and current records of ports that meet the above conditions, perform the following steps:

Read NetFlow data and process log data; Computing principles from historical data sources; Determine the importance of state value; Identify the host epidemic. The specific process form is described below.

(1) *Select a baseline to analyze NetFlow records and perform statistics.* From the data flow collected by NetFlow, locate the destination destination port (dst port) and identify it to determine whether it is the destination destination destination for the specified data [11]. If it is a known port, first check whether the target port is a blacklist port. If so, calculate the traffic data of the port and collect them into historical data for future calculation; If it is an unknown port, first check whether the destination port belongs to the whitelist port. If it is not, then count the traffic records of all relevant ports as a whole (with port number -1 indicating these overall ports). Network avoidance can include three main types: victim avoidance, lightweight avoidance, and performance avoidance. Most types of abuse in enterprises are abnormal, as shown in Table 2.2.

Table 2.2 shows the relationship between different traffic and the number of packets or bytes. Taking a worm as an example, there will be more data packets or bytes in terms of network traffic. However, the number of bytes transmitted by the worm is much smaller than that of traditional data, indicating that during transmission of the worm, the number of data packets in the network will increase, while byte statistics will not change much [12]. Therefore, at this point, counting the number of packets according to the response rate is more effective, which can detect the presence of abnormal worms and other network traffic. On the contrary, for flash memory congestion, such as some downloading attributes, the number of packets or bytes will increase, but the impact of downloading traffic costs on network performance is more significant. the effect of network performance is more significant. Therefore, in this case, the value of the byte count can have a better effect on whether the network is uniquely or unequally distributed.

(2) *Calculate baseline values from historical table data.* If the members of a company or an organization have a similar behavior, their network traffic still has a close relationship, that is, due to the work and rest of the members of the organization, the network traffic has a consistent pattern and the peak release status. According

to this characteristic, we divide the algorithm into two types: working day and non-working day. Week: Within 5 minutes, take the sample points of the corresponding port for the current minute of the previous 20 working days, calculate their average and standard deviations, and use the Grubbs method to eliminate negative results. Non working days: In units of 5 minutes, take a port sampling point corresponding to the current minute of the first 8 non working days, remove the maximum and minimum 5% data, and then take the average to obtain the current hourly baseline. Taking workdays as an example, the specific steps are described as follows:

Sort the flow values of the device in the first 20 working days for each port that meets the criteria, and obtain the following sequence:

$$\{X_{1_i}, X_{2_i}, \dots, X_{20_i}, i_j \in \{1, 2, \dots, 20\} X_{1_i} \leq X_{2_i} \dots \leq X_{20_i}\} \quad (2.1)$$

Remove the maximum and minimum values by 5%, leaving the following sequence:

$$\{X_{1_i}, X_{2_i}, \dots, X_{19_i}, i_j \in \{1, 2, \dots, 20\} X_{1_i} \leq X_{2_i} \dots \leq X_{19_i}\} \quad (2.2)$$

Find the average value, which is the baseline value corresponding to a port that meets the conditions during that period.

(3) *Determine dynamic critical state values.* Set  $T_h$  as the baseline threshold for a time period subdivision, and  $b_1$  as the baseline value. The larger the value of  $b_1$ , the greater  $T_h$  should be, that is,  $b_1$  is proportional to  $T_h$ . There are:

$$T_h = K \times b_1 \quad (2.3)$$

Assuming  $K$  is a constant greater than 1, make the necessary modifications to  $K$  value based on historical data. Simple analysis shows that the higher the  $K$  value, the easier it is to determine the transmission behavior, but it can also lead to an increase in detection rate, making some less parasitic bacteria can be detected [13]; On the contrary, the smaller the  $K$  value, the more convenient the transmission behavior, which can also increase the false value. Some normal information flows will also be misjudged as abnormal traffic. The system established by the author initially sets the  $K$  value to 2, meaning that abnormal alarms are only detected when the current value is twice the baseline value.

(4) *Identify the source of abnormal traffic.* When the traffic is monitored more than the critical value described in step 3 in a certain period of time, this cycle is called an abnormal cycle. Track the data during this period for further analysis and identification of host generated traffic defects based on NetFlow data. The key step is to identify raw data on NetFlow, sort the count byte or count packet at that time, port, or port from the top, and identify the master IP address and port for further processing.

#### 2.4. Real time performance monitoring related technologies.

(1) *Basic Principles.* Real time performance monitoring technology includes the following categories: Focusing on performance indicator monitoring, comparing with indicator thresholds, historical statistical baselines, etc., to discover the degradation of business and network performance; Associate alarms with engineering cut-over and configuration information, eliminate abnormal alarms, reduce invalid alarm rates, and improve work efficiency; Assisted by business dial testing and signaling detection to locate or discover faults; Real time performance monitoring technology focuses on discovering network anomalies and alarm generation mechanisms, but does not elaborate on the processing measures to be taken after the alarm is generated.

(2) *Indicator selection.* Real time performance monitoring captures the overall operational status of the business, promptly detects business or network anomalies, and monitors network operational status from the perspective of user business. Correspondingly, the selection of real-time performance monitoring performance indicators should reflect the operational status of the business or network. A feasible method for selecting performance indicators is to analyze business processes, decompose and refine them layer by layer from the perspective of business processes, and extract performance indicators in business process analysis. The selection of real-time performance monitoring indicators should follow the following principles and basic methods: Following the logical path of "Purpose>Method>Indicator Set", starting from analyzing the correspondence between business processes and network element devices, select complete indicators that match the positioning. Starting from customer perception, propose performance indicators that reflect the "end-to-end" business quality;

Propose supplementary performance alert indicators to address the limitations of network element capabilities and insufficient network management functions; Starting from the purpose of reducing the occurrence rate of faults, select performance indicators that can timely reveal the hidden dangers of faults, such as statistical channel utilization indicators within the base station range, in order to explain channel issues, if the scope is expanded to BSC, it is possible that a channel issue cannot affect the entire indicator, leading to valuable information being hidden; Meeting the principle of minimizing the set of indicators, that is, when multiple indicators can reflect the same fault or business degradation, only the indicators that can directly reflect the problem are selected, and other indicators are only queried and called during in-depth analysis, based on the high incidence of faults and complaints about network elements, select the key points. According to the fact that 75% of the complaints on the GSM network side belong to the wireless part, the traffic network will focus on the wireless part; The sources of various indicators need to be clearly defined (such as existing network management, Yu Ling method, automatic reporting of network elements, business call testing, etc.), collection granularity, latency requirements, feasibility, and the impact of monitoring work on network elements needs to be evaluated.

(3) *Relevant regulations on indicator data.* Indicator data source: From an analysis perspective, most performance indicators have been collected and stored in the network management system, and meet the time and spatial granularity requirements for indicator collection. These indicators can be directly extracted based on the existing network management system. For indicators that have been collected and stored in the network management system, but do not meet the time or spatial granularity requirements for indicator collection, or that cannot be provided by the existing network management system yet, this can mainly be achieved by modifying the network management system, if the network element's capabilities permit, increase the collection time density and refine the spatial granularity of the collection. In the case where the network management system renovation cannot meet the real-time performance monitoring requirements, it can be considered to achieve this by directly connecting to the network element, using instruction interaction, collecting device operation logs, and logging in to the device to obtain statistical reports. Data collection principle: For situations where instructions are required to retrieve network element data, the impact of the instructions on the network element itself should be considered. It is recommended to set the switch (protection threshold) for obtaining data in the real-time performance monitoring function module or network management system. When the system load is too high, the real-time performance monitoring module will automatically shut down the collection of this indicator. After the system load is normal, collection can be resumed to avoid further aggravation of the system load.

The principle for setting the granularity of collection time: Time granularity is the minimum period for extracting indicator data. The set value should be suitable for the monitoring needs. If it is too small, it will cause excessive system load pressure but not generate actual value; Excessive size may lead to loss of monitoring significance [1]. At present, the business quality data provision capabilities of mainstream devices include different granularity such as 1 minute, 5 minutes, 15 minutes, and 1 hour. The granularity for selecting business quality indicators for real-time performance monitoring is generally 15 minutes. In special cases, finer or more relevant granularity can be considered. The system connection rate indicator of MSC is a comprehensive indicator that can comprehensively reflect network issues, covering the overall performance of wireless and switching systems. The user perception is obvious, and considering the impact of collection on the system, it is recommended to use a time granularity of 15 minutes. The location update success rate of HLR reflects the successful login of users to the network, which can reflect the use of illegal cards, SIM card errors, office data errors, etc. Considering the impact of its collection on business systems, it is recommended to use a time granularity of 1 hour. The principle for determining the collection delay requirement: The delay of business quality alarms is directly related to the size of real-time performance monitoring capabilities. However, business quality alerts require a complete process of data generation, extraction, analysis, and presentation. From the perspective of balancing real-time requirements and feasibility, the delay cannot exceed one data cycle. For data with a granularity of 15 minutes, the delay should not exceed 15 minutes. The system presentation must be completed before 10:30 from 10:00 to 10:15.

**3. Result analysis.** This experiment mainly collects Net Flow data from the backbone router of a certain power enterprise information network, and uses running Oracle stored procedures to calculate baseline values

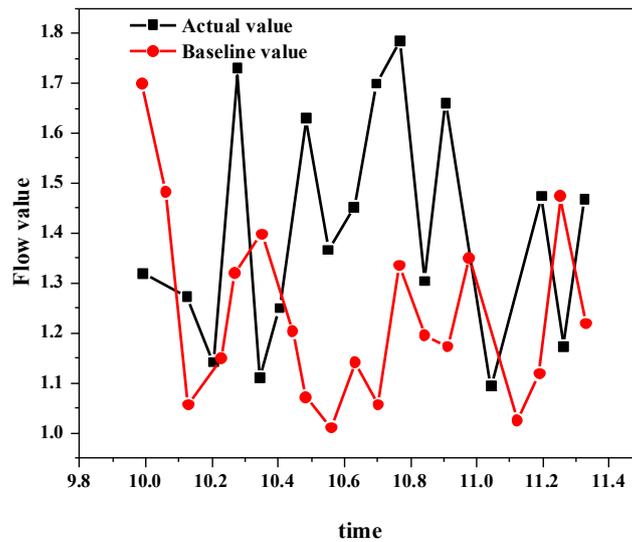


Fig. 3.1: Unknown port traffic value 1

for result analysis. Taking the data results from working days as an explanatory analysis, the situation is similar for non working days. Figure 3.1 shows the unnamed port traffic values for a period of real traffic and calculated baseline values from 10:00 to 11:35 on a certain working day, which is a normal situation and there is no alarm situation.

Figure 3.2 shows the unnamed port traffic values for a period of real traffic and calculated baseline values from 14:00 to 15:35 on a certain working day. The actual value exceeds the baseline value by more than twice, which is an abnormal situation and requires an alarm. The IP and port that generated large traffic need to be analyzed through the original NetFlow record, and the reason should be analyzed.

Figure 3.3 shows the actual traffic and calculated baseline values for TCP port 135 on a certain working day from 16:00 to 17:35. This port is used for shock waves, with a byte count of 48 [7]. From the graph, we can see that at 16:45, the traffic value of Port 135 suddenly increased, and then returned to normal. It may be that a certain machine was infected with the shockwave virus. When the machine tried to infect other machines, it was blocked by a firewall, causing the traffic value to return to normal.

**4. Conclusion.** A baseline algorithm based on the method for detecting traffic anomalies in enterprises is proposed, and a baseline algorithm with multiple time-frequency segments and multiple communication ports is proposed. The baseline algorithm is proposed. Based on this dynamic principle, the analysis is made on the traffic data of operation and failure date in order to achieve the purpose of detecting the traffic malfunction. Based on the experimental results, the advantages of this method are as follows: Adopting baseline analysis method has the advantage of singularity. The baseline based abnormal traffic monitoring method proposed by the author has a high degree of singularity, which is analyzed in multiple time periods of working and non working days. Only the baseline is used as the sole criterion for analysis in monitoring, avoiding the tedious and complex monitoring methods of other methods. Different detection methods will not be used according to the reasons for abnormal traffic generation, this overall mechanism presents a high degree of singularity in the analysis method. The monitoring method proposed by the author, although lacking in accuracy and unable to fully monitor abnormal traffic, has been greatly simplified due to its highly singular mechanism. It is very useful for a complex and large network of a large enterprise, with improved efficiency and better real-time performance. However, the first base algorithm proposed by the author is simple, and the importance of multi-bases has a significant impact on the accuracy of the algorithm. In the future, it is necessary to improve the basis generating algorithm while control complexity to realize traffic congestion.

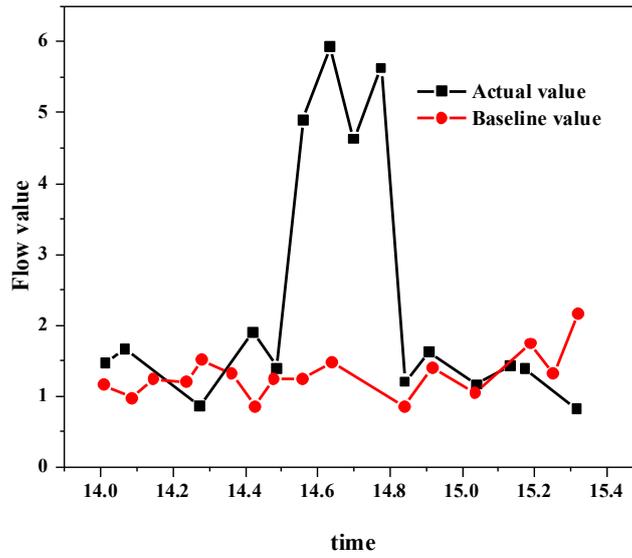


Fig. 3.2: Unknown port traffic value 2

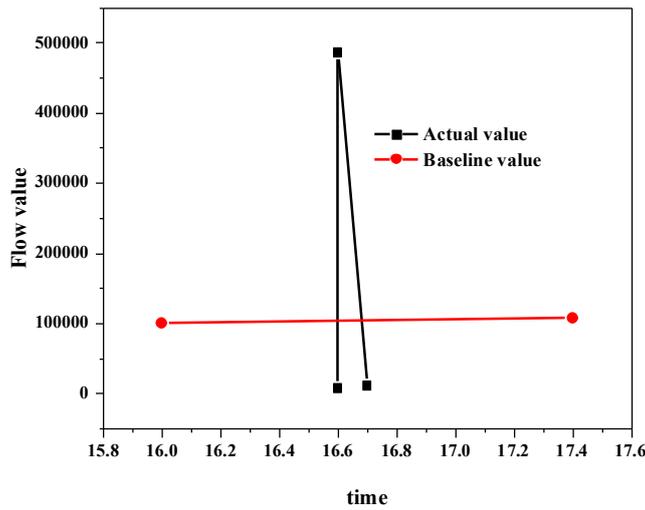


Fig. 3.3: Flow Value of Famous Port (135 Port)

**Acknowledgement.** Jiangsu Collaborative Innovation Center for Building Energy Saving and Construction Technology:XT-Research and application of public building energy consumption monitoring system based on big data, Project No: SJXTY1603, Project Leader: Zhaoli Wu.

REFERENCES

[1] J. AZIMJONOV, A. ÖZMEN, AND M. VARAN, *A vision-based real-time traffic flow monitoring system for road intersections*, Multimedia Tools and Applications, (2023), pp. 1–20.  
 [2] C. A. BOLLMANN, M. TUMMALA, AND J. C. MCEACHEN, *Resilient real-time network anomaly detection using novel non-parametric statistical tests*, Computers & Security, 102 (2021), p. 102146.  
 [3] L. DUAN, J. ZHOU, Y. WU, AND W. XU, *A novel and highly efficient botnet detection algorithm based on network traffic*

- analysis of smart systems*, International Journal of Distributed Sensor Networks, 18 (2022), pp. 182459–182476.
- [4] Q. GONG, P. DEMAR, AND M. ALTUNAY, *Thundersecure: deploying real-time intrusion detection for 100g research networks by leveraging stream-based features and one-class classification network*, International Journal of Information Security, 21 (2022), pp. 799–812.
  - [5] J. GUO, X. DING, AND W. WU, *Reliable traffic monitoring mechanisms based on blockchain in vehicular networks*, IEEE Transactions on Reliability, 71 (2021), pp. 1219–1229.
  - [6] S. HE, L. CHEN, S. ZHANG, Z. GUO, P. SUN, H. LIU, AND H. LIU, *Automatic recognition of traffic signs based on visual inspection*, IEEE Access, 9 (2021), pp. 43253–43261.
  - [7] A. M. HTUT AND C. ASWAKUL, *Development of near real-time wireless image sequence streaming cloud using apache kafka for road traffic monitoring application*, PLoS One, 17 (2022), p. e0264923.
  - [8] J. JIAO, X. SUN, Y. ZHANG, L. LIU, J. SHAO, J. LYU, AND L. FANG, *Modulation recognition of radio signals based on edge computing and convolutional neural network*, Journal of Communications and Information Networks, 6 (2021), pp. 280–300.
  - [9] A. KRISHNAKUMAR, S. KADIAN, U. HEREDIA RIVERA, S. CHITTIBOYINA, S. A. LELIÈVRE, AND R. RAHIMI, *Organ-on-a-chip platform with an integrated screen-printed electrode array for real-time monitoring trans-epithelial barrier and bubble formation*, ACS Biomaterials Science & Engineering, 9 (2023), pp. 1620–1628.
  - [10] C. LI, Q. PENG, D. WANG, L. LUO, AND H. ZUO, *Smart substation network quality monitoring and fault prediction*, 2108 (2021), p. 012061.
  - [11] D. LI, J. BAO, S. YUAN, H. WANG, L. WANG, AND W. LIU, *Image enhancement algorithm based on depth difference and illumination adjustment*, Scientific Programming, 2021 (2021), pp. 1–10.
  - [12] X. LIU, Q. LI, W. CHEN, P. SHEN, Y. SUN, Q. CHEN, J. WU, J. ZHANG, P. LU, H. LIN, ET AL., *A dynamic risk-based early warning monitoring system for population-based management of cardiovascular disease*, Fundamental Research, 1 (2021), pp. 534–542.
  - [13] ———, *A dynamic risk-based early warning monitoring system for population-based management of cardiovascular disease*, Fundamental Research, 1 (2021), pp. 534–542.
  - [14] J. PANG AND Z. ZHAO, *Real-time monitoring of fluidized bed agglomerating based on improved adaboost algorithm*, 1924 (2021), p. 012026.
  - [15] X. QI, Y. JI, W. LI, AND S. ZHANG, *Vehicle trajectory reconstruction on urban traffic network using automatic license plate recognition data*, IEEE Access, 9 (2021), pp. 49110–49120.
  - [16] F. I. H. SAKIYAMA, F. LEHMANN, AND H. GARRECHT, *A novel runtime algorithm for the real-time analysis and detection of unexpected changes in a real-size shm network with quasi-distributed fbg sensors*, Sensors, 21 (2021), p. 2871.
  - [17] Y. WANG, Q. WANG, D. SUO, AND T. WANG, *Retraction note: Intelligent traffic monitoring and traffic diagnosis analysis based on neural network algorithm*, Neural Computing and Applications, 35 (2023), pp. 4183–4183.
  - [18] H. WU, Y.-P. ZHAO, T.-L. YANG, AND H.-J. TAN, *An ensemble radius basis function network based on dynamic time warping for real-time monitoring of supersonic inlet flow patterns*, Aerospace Science and Technology, 111 (2021), p. 106551.
  - [19] X. XIANG, C. CHEN, H. WANG, H. LU, H. ZHANG, AND J. CHEN, *A real-time processing method for gb-sar monitoring data by using the dynamic kalman filter based on the ps network*, Landslides, (2023), pp. 1–17.
  - [20] W. YU, K. RUAN, H. TANG, AND J. HUANG, *Routing hypergraph convolutional recurrent network for network traffic prediction*, Applied Intelligence, 53 (2023), pp. 16126–16137.

*Edited by:* B. Nagaraj M.E.

*Special issue on:* Deep Learning-Based Advanced Research Trends in Scalable Computing

*Received:* Sep 14, 2023

*Accepted:* Nov 6, 2023



## APPLICATION OF IMPROVED PSO AND BP HYBRID OPTIMIZATION ALGORITHM IN ELECTRICAL AUTOMATION INTELLIGENT CONTROL SYSTEM

LIJING LI\*, XIAOJIAN WANG<sup>†</sup> AND MEI YANG<sup>‡</sup>

**Abstract.** A fuzzy RBF-PID control strategy based on particle swarm optimization (PSO) algorithm is proposed to solve the problem of large inertia lag in temperature control system of industrial production refuse furnace. In this control system, an improved particle swarm optimization algorithm combined with inertia weight and genetic transformation was used to optimize the initial values of membership functions of fuzzy RBF (radial basis function). Then, BP (error backpropagation) algorithm is used for fine tuning, and fuzzy reasoning and RBF learning ability are combined to adjust the PID control parameters online to achieve the optimal PID control effect. The simulation results show that the algorithm has fast tracking, small overshoot, and is not easily trapped in local minima. At the same time, its robustness and anti-interference performance are better than traditional PID control.

**Key words:** Improving PSO, BP hybrid optimization algorithm, Electrical automation, Intelligent control system

**1. Introduction.** PSO algorithm is the most popular optimization algorithm in recent years. Combining PSO algorithm with BP neural network algorithm can solve the shortage of BP neural network itself. Using particle swarm optimization algorithm instead of gradient descent in BP network can optimize the link weight of each layer of BP network, thus improve its generalization ability and learning ability, and greatly improve its convergence speed [12]. In the process of optimizing BP neural network using particle swarm optimization algorithm, the iterative method of particle swarm optimization is usually used to replace the gradient improvement in BP neural network algorithm, and PSO algorithm is used to optimize the training weight of BP neural network. The key to using PSO algorithm to solve the problems in BP neural networks is to achieve the following two points: the particle dimension value in particle swarm optimization algorithm corresponds to the number of weighted combinations in BP neural network. Using the size of each particle in PSO according to the weight of each component in BP neural network. The key to the PSO algorithm in pronunciation, based on multi-pronunciation vector and the number of weights as equal to the size of pronunciation. Replaces the motor of the particle swarm algorithm with the mean square error function in BP neural network, and uses the energy potential to research the ability of the particle swarm algorithm to reduce the mean square error of BP network. Resistance furnace is an electrical heating equipment commonly used in fields such as machinery, chemical engineering, and manufacturing. At present, the furnace temperature control system generally adopts conventional PID control, which cannot achieve good control results for objects that are difficult to establish accurate models. Therefore, it is urgent to find a fast and accurate temperature control system. With the development of science and technology, PID control and some intelligent control have been combined and applied in industrial production. For this research question. Song, W. et al The macro and micro models were established to predict the maintenance activities of the combat troops, and the parameters were analyzed. A prediction model for equipment maintenance activities that meet the requirement of real-time mobile echelon storage was established. In the condition of meeting annual motor stock balance, an improved particle swarm optimization hybrid optimization algorithm is developed to solve the model based on the time driving consumption from the equipment allocated to the annual training program. the particle swarm optimization hybrid optimization algorithm is proposed. An example study was made on a few cars in some room [14]. In order to solve the above problems, a fuzzy RBF-PID control method based on improved PSO algorithm is adopted. This method first solves the problem of local extremum and early convergence in particle swarm optimization algorithms. By

---

\*Shijiazhuang Information Engineering Vocational College, Shijiazhuang, Hebei, 052161, China.

<sup>†</sup>Shijiazhuang Information Engineering Vocational College, Shijiazhuang, Hebei, 052161, China. (Corresponding Author)

<sup>‡</sup>Shijiazhuang Information Engineering Vocational College, Shijiazhuang, Hebei, 052161, China.

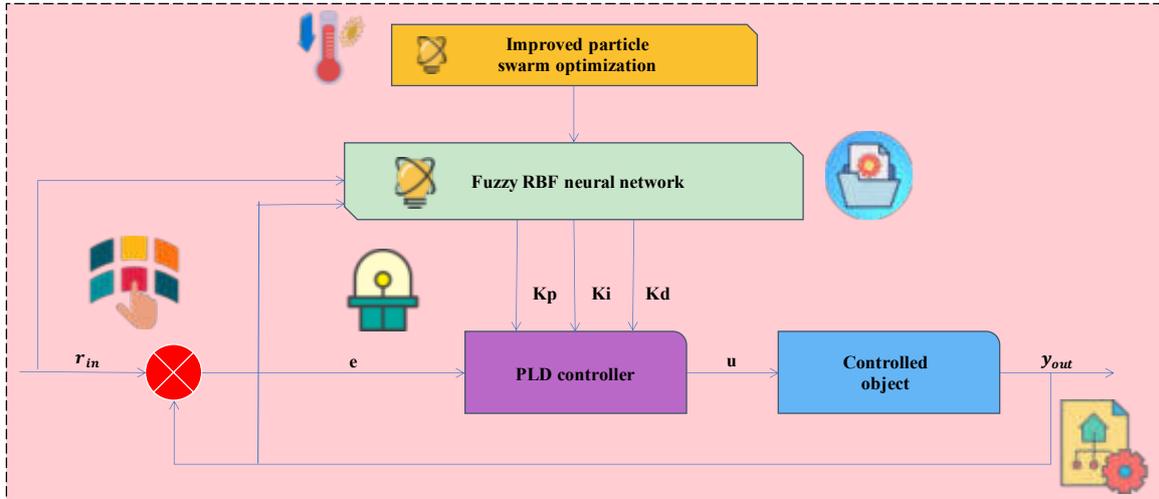


Fig. 2.1: Structure diagram of temperature control system

adding the modified algorithm, an improved PSO algorithm was established, and the primary importance of fuzzy RBF neural network was optimized using this strategy. Then, BP (Error Backpropagation) algorithm is used to adjust the key, center, and width of RBF neural network. Finally, combined with the power expression ability of fuzzy logic and self-learning ability of RBF neural network, PID is not adjusted online. Matlab test results show that this method has good control, adaptability, and robustness [5, 11].

## 2. Methods.

**2.1. Resistance furnace temperature control system.** The resistance furnace is a complex controlled object with large hysteresis and inertia. Due to the inconvenient modification of control parameters in traditional PID controllers and the difficulty in ensuring their control quality in terms of control accuracy, stability time, and overshoot, the author proposes to apply fuzzy RBF tuning PID control technology to the furnace temperature control system for resistance furnace temperature. The structure of its temperature control system is shown in Figure 2.1, where PSO is improved to optimize the weights of the fuzzy RBF neural network. The system's given value  $r_{in}$  and the controlled object's output  $y_{out}$  are combined with the fuzzy RBF neural network and the three parameters ( $K_p, K_i, K_d$ ) of the PID controller to form a functional relationship, the controller to output  $u$ , and then combined with the controlled object output  $y_{out}$  to form a complete closed-loop control structure [8, 4].

### 2.2. Improving PSO and BP hybrid optimization algorithm.

(1) *Hybrid optimization algorithm.* In the PID control of fuzzy RBF neural network, the selection of parameters  $w_{ij}$ ,  $c_{ij}$ , and  $b_{ij}$  has a significant impact on the approximation ability of the network function, the author selects a hybrid optimization algorithm that combines improved PSO and BP with inertia weight factor and genetic mutation operator to adjust its parameters. Firstly, PSO with inertia weight has been applied to international research. If an earlier result occurs, a change is made to take the command to jump out of the best locale and search for objects in other locales until a global solution is found. Then, an adaptive algorithm is designed using BP (Error Backpropagation) algorithm to get the optimal control parameters [7, 9].

(2) *Improved particle swarm optimization algorithm.* In order to expand particle space exploration and improve global and local searching ability, inertia weight  $w$  is introduced into the PSO model, and its velocity and position iteration method are as follows:

$$v_{id}(t+1) = wv_{id}(t) + c_1r_1(p_{id}(t) - x_{id}(t)) + c_2r_2(p_{gd}(t) - x_{id}(t)) \quad (2.1)$$

$$x_{id}(t+1) = x_{id}(t) + v_{id}(t+1) \quad (2.2)$$

$$w(t) = \begin{cases} 1 \times \frac{t}{G} + 0.4, & 0 \leq \frac{t}{G} \leq 0.5 \\ -1 \times \frac{t}{G} + 1.4, & 0.5 \leq \frac{t}{G} \leq 1 \end{cases} \tag{2.3}$$

In the formula,  $c_1$  and  $c_2$  are learning factors, and  $c_1 = c_2 = 2$ ;  $t$  is the number of iterations;  $1 \leq i \leq m, 1 \leq d \leq N$ ;  $r_1$  and  $r_2$  take the random number between (0,1);  $G$  is the total number of iterations [10].

In general, particles update their state at the next moment based on their current position and speed. However, if the target of the particle search has already fallen into the local extreme, the particles are limited to the local space and cannot be fully optimized in the entire search space. In order to enable the particles to jump out of the local optimum and discover the global optimal position, a mutation operator is introduced to mutate the particles that have fallen into the local extreme with probability  $p$ . The formula is as follows:

$$p = \begin{cases} k, \sigma^2 < \sigma_d, J(P_{gbest}) > J_d \\ 0 \end{cases} \tag{2.4}$$

In the equation,  $k$  is [0.1, 0.3] Any value between;  $\sigma^2$  is the variance of population fitness; The value of  $\sigma_d^2$  is generally much smaller than the maximum value of  $\sigma^2$ ;  $J_d$  is set to the theoretical optimal value [17]. The author employs a random perturbation approach to perform mutation operations on the global extreme value ( $p_{gbest}$ ), where  $p_{gbest,k}$  is the  $k$ -th dimensional value of  $p_{gbest}$ ,  $\eta$  Following a normal (0, 1) distribution, that is:

$$p_{gbest,k} = p_{gbest,k}(1 + 0.5\eta) \tag{2.5}$$

The improved methods of PSO algorithm are as follows: ① initial total size  $m$  of particle swarm, the velocity  $c_1$  and  $c_2$ , the number of iterations  $G$ , and velocity and position [18]. ② Select the appropriate fitness  $J$  and use ITAE (the wrong rate of time error) criterion as the fitness function for improving PSO algorithm. The calculation formula is as follows:

$$J = \int_0^{+\infty} t|e(t)|dt \tag{2.6}$$

(3). Find the fitness values of each particle, sort them according to the minimum value, and obtain the particles corresponding to the minimum value as the initial position of the global extremum. At the same time, the initial position corresponding to the particle fitness value is the initial position of the local extremum [20, 3]. ④ Calculate the inertia weight factor  $w$  according to equation (2.3), then calculate the mutation probability  $p$  according to equation (2.4), and randomly select the random number  $n$  between (0, 1). If  $n < p$ , then perform the mutation operation according to equation (2.5). ⑤ Based on models (2.1) and (2.2), the velocity and position of particles in the flame are updated. ⑥ Calculate new exercise results, and then update the effectiveness of individuals and groups. When the number of iterations reaches the preset number, stop the iteration, otherwise it will turn to (2).

**3. Result Analysis.** The author designed the controller using a resistance furnace as the controlled object, and the transfer function of the controlled object is:

$$G(s) = \frac{K}{Ts + 1} e^{-Ts} = \frac{1}{60s + 1} e^{-80s} \tag{3.1}$$

This system uses the unit step signal as the input signal, utilizes practical experience and three other control parameters, and then combines the author's control system. After multiple simulation experiments, as shown in Table 3.1, the final system parameters are set as follows: Sampling time of 20 seconds, learning  $\eta = 0.15$ , inertia coefficient  $a=0.02$ , the initial value ranges of the center vector  $c_{ij}$  and the base width vector  $b_{ij}$  of the Gaussian function are  $[-6, 6], [0.1, 3]$ , respectively, the initial value range of network weight  $w_{ij}$  is  $[-1, 1]$ , the total number of population particles is 60, the total number of iterations is 300, and the dimension of particles is  $N = 168$ , which means there are a total of 168 parameters that need to be optimized [16].

Figures 3.1 and 3.2 show the simulation comparison and fitness function  $J$  optimization process without input disturbance, while Figures 3.3 and 3.4 show the simulation comparison and fitness function  $J$  optimization process with 10% disturbance added to the system at 600 seconds, from the figure, it can be seen that compared

Table 3.1: Parameter simulation experiment

Learning Rate	inertial coefficient	Base width vector	center vector	Network weight	Is it stable	overshoot	Time when $e$ is zero
0.16	0.02	[0.1 3]	[-3 3]	[-1 1]	Yes	7.41	125.1
0.21	0.02	[0.1 3]	[-3 3]	[-1 1]	Yes	12.21	137.1
0.16	0.02	[0.1 3]	[-4 4]	[-1 1]	No	-	-
0.21	0.02	[0.1 3]	[-4 4]	[-1 1]	No	-	-
0.16	0.02	[0 4]	[-3 3]	[-1 1]	No	-	-
0.21	0.02	[0 4]	[-3 3]	[-1 1]	No	-	-
0.16	0.02	[0 4]	[-6 6]	[-1 1]	Yes	-	-
0.21	0.02	[0 4]	[-6 6]	[-1 1]	No	-	-
0.11	0.02	[0.1 3]	[-6 6]	[-1 1]	Yes	not have	300.1
0.16	0.02	[0.1 3]	[-6 6]	[-1 1]	Yes	not have	120.1
0.21	0.02	[0.1 3]	[-6 6]	[-1 1]	Yes	2.81	151.1
0.26	0.02	[0.1 3]	[-6 6]	[-1 1]	Yes	9.81	293.2

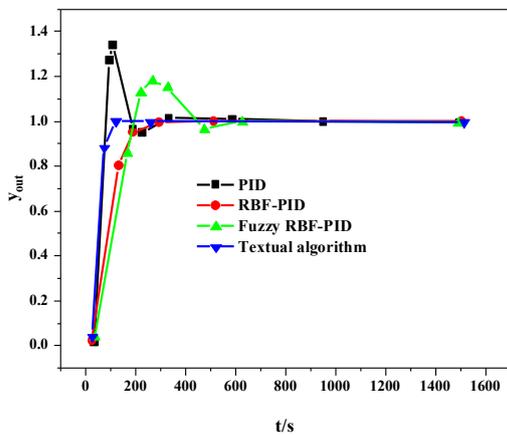


Fig. 3.1: Comparison of simulation curves

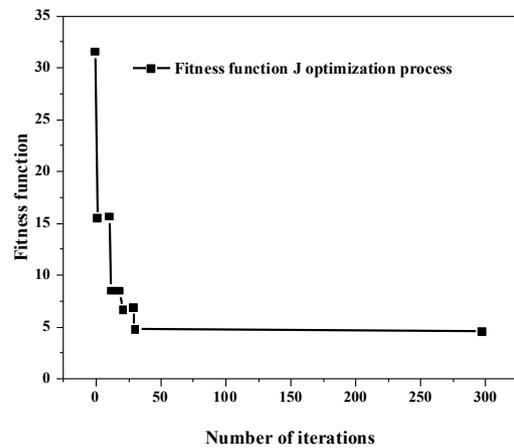


Fig. 3.2: Optimization process of fitness function  $J$

with traditional PID controllers, RBF-PID controllers have a reduced overshoot and are prone to oscillations, making it difficult for the system to achieve rapid stability. However, adopting fuzzy RBF-PID control strategy can significantly improve system performance, but the adjustment time is long [13, 2, 6, 19, 15, 1]. In comparison, the fuzzy RBF-PID control method using improved particle swarm optimization algorithm has short response time, good anti-interference performance, and demonstrated good control accuracy.

**4. Conclusion.** The author proposes a fuzzy RBF-PID control strategy based on improved particle swarm optimization algorithm for resistance furnace temperature control in industrial production, which addresses the issues of large inertia and large lag of temperature control objects. This algorithm not only combines the characteristics of traditional PID controllers, but also has reasoning and self-learning abilities, overcoming the problems of traditional PID controllers being unable to adjust parameters online and having poor adaptability. At the same time, the algorithm introduces inertia weight and personnel variables to improve the searching ability of particle swarm, avoids the early capture of local extremum by particles, and uses ITAE objective function as the driving mechanism of the improved PSO algorithm. Through Matlab simulation analysis, the performance of the fuzzy RBF-PID control algorithm based on improved PSO algorithm is superior to other

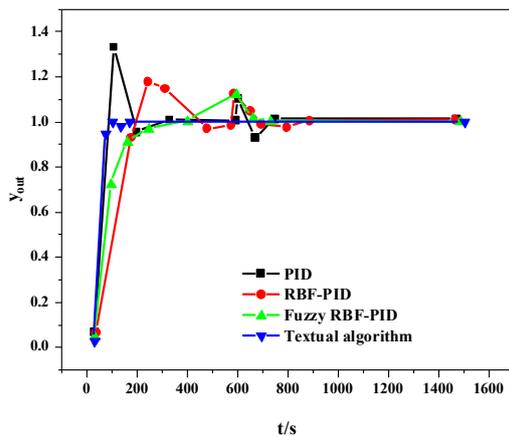


Fig. 3.3: Comparison of simulation curves with 10% disturbance added to the system

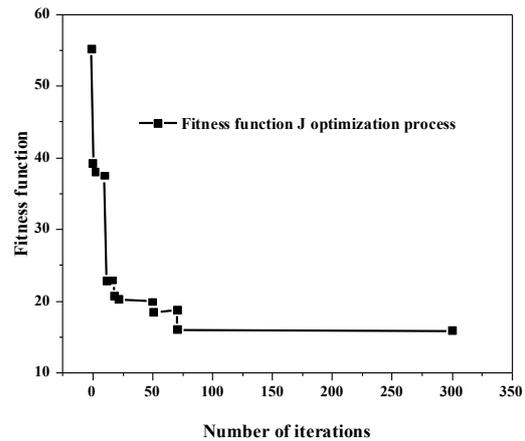


Fig. 3.4: Optimization process of fitness function J with 10% disturbance in the system

PID control algorithms, with shorter response time and better convergence.

#### REFERENCES

- [1] K. CHEN, J. YU, M. HAO, C. ZHANG, B. YANG, H. TAN, AND Y. WANG, *Research on intelligent blending method of matcha based on improved pid control algorithm*, Journal of Physics: Conference Series, 1848 (2021), p. 012014.
- [2] H. FENG, W. MA, C. YIN, AND D. CAO, *Trajectory control of electro-hydraulic position servo system using improved pso-pid controller*, Automation in Construction, 127 (2021), p. 103722.
- [3] Y. FENG, M. WU, L. CHEN, X. CHEN, W. CAO, S. DU, AND W. PEDRYCZ, *Hybrid intelligent control based on condition identification for combustion process in heating furnace of compact strip production*, IEEE Transactions on Industrial Electronics, 69 (2021), pp. 2790–2800.
- [4] G. HAFEZ, I. KHAN, S. JAN, I. A. SHAH, F. A. KHAN, AND A. DERHAB, *A novel hybrid load forecasting framework with intelligent feature engineering and optimization algorithm in smart grid*, Applied Energy, 299 (2021), p. 117178.
- [5] Y. HAN, C. LIU, H. XIU, Z. LI, S. SHAN, X. WANG, L. REN, AND L. REN, *Trajectory control of an active and passive hybrid hydraulic ankle prosthesis using an improved pso-pid controller*, Journal of Intelligent & Robotic Systems, 105 (2022), pp. 1–20.
- [6] W. HONG, I. CHAKRABORTY, H. WANG, AND T. GANG, *Co-optimization scheme for hybrid electric vehicles powertrain and exhaust emission control system using future speed prediction*, IEEE Transactions on Intelligent Vehicles, 6 (2021).
- [7] R. KUMAR, S. DIWANIA, P. KHETRAPAL, AND S. SINGH, *Performance assessment of the two metaheuristic techniques and their hybrid for power system stability enhancement with pv-statcom*, Neural Computing and Applications, 35 (2022), pp. 3723–3744.
- [8] K. LIU, H. ZHANG, B. ZHANG, AND Q. LIU, *Hybrid optimization algorithm based on neural networks and its application in wavefront shaping*, Optics Express, 29 (2021), pp. 15517–15527.
- [9] P. LIU, Q. HU, K. JIN, G. YU, AND Z. TANG, *Toward the energy-saving optimization of wlan deployment in real 3-d environment: A hybrid swarm intelligent method*, IEEE Systems Journal, 16 (2021), pp. 2425–2436.
- [10] S. MAIHEMUTI, W. WANG, H. WANG, AND J. WU, *Voltage security operation region calculation based on improved particle swarm optimization and recursive least square hybrid algorithm*, Journal of Modern Power Systems and Clean Energy, 9 (2020), pp. 138–147.
- [11] S. NASROLLAHZADEH, M. MAADANI, AND M. A. POURMINA, *Optimal motion sensor placement in smart homes and intelligent environments using a hybrid woa-pso algorithm*, Journal of Reliable Intelligent Environments, 8 (2022), pp. 345–357.
- [12] X. NIE AND J. LUO, *The hybrid intelligent optimization algorithm and multi-objective optimization based on big data*, Journal of Physics: Conference Series, 1757 (2021), p. 012132.
- [13] K. RAJESH, *Radial basis function neural network mppt controller-based microgrid for hybrid stand-alone energy system used for irrigation*, Circuit World, 49 (2023), pp. 251–266.
- [14] W. SONG, J. WU, H. CHANG, AND G. XU, *Equipment maintenance task prediction model analysis based on bas-pso hybrid optimization algorithm*, Systems Science & Control Engineering, 9 (2021), pp. 455–466.
- [15] M. SUHAIL, I. AKHTAR, S. KIRMANI, AND M. JAMEEL, *Development of progressive fuzzy logic and anfis control for energy management of plug-in hybrid electric vehicle*, IEEE Access, 9 (2021), pp. 62219–62231.
- [16] Q. SUN, J. YANG, S. MA, Y. HUANG, Y. YUAN, AND Y. HOU, *3d vessel extraction using a scale-adaptive hybrid parametric tracker*, Medical & Biological Engineering & Computing, 61 (2023), pp. 2467–2480.

- [17] Z. WANG, S. DENG, P. DUAN, Y. XU, AND W. LI, *Towards speech recognition and training utilization in the nuclear power main control room*, Journal of Physics: Conference Series, 2242 (2022), p. 012028.
- [18] Y. YAN, Y.-X. QIN, L.-S. ZHANG, T. JIA, AND F. SUN, *Swing suppression control in quayside crane by using fuzzy logic and improved particle swarm optimization algorithm*, Iranian Journal of Science and Technology, Transactions of Mechanical Engineering, 47 (2023), pp. 1131–1144.
- [19] J. ZHANG, J. ZHANG, F. ZHANG, M. CHI, AND L. WAN, *An improved symbiosis particle swarm optimization for solving economic load dispatch problem*, Journal of Electrical and Computer Engineering, 2021 (2021), pp. 1–11.
- [20] Y. ZHAO, J. CHEN, AND W. LI, *Improved whale algorithm and its application in cobot excitation trajectory optimization*, International Journal of Intelligent Robotics and Applications, 6 (2022), pp. 615–624.

*Edited by:* B. Nagaraj M.E.

*Special issue on:* Deep Learning-Based Advanced Research Trends in Scalable Computing

*Received:* Sep 18, 2023

*Accepted:* Nov 7, 2023



## DESIGN OF COMPUTER INFORMATION MANAGEMENT SYSTEM BASED ON MACHINE LEARNING ALGORITHMS

YAN LI\*

**Abstract.** In order to improve the efficiency of office automation, regulate work frequency, and improve office efficiency, this paper presents a computer information management system design based on machine learning technology. Firstly, the basic design principles of computer information management systems are analyzed, and secondly, risk prediction is studied. The risk of computer information management systems is caused by the cross influence of different risk factor indicators, and has linear and nonlinear characteristics. Using a single prediction model cannot obtain accurate prediction results. Therefore, the risk prediction method for computer information management based on machine learning technology. The risk prediction method is established by using Analytic Hierarchy Method in machine learning algorithms, and the historical data is collected according to the index system. The weight of the initial prediction is determined by the combination of subjective and objective weight; In machine learning algorithms, risk prediction and benefit prediction are used as input and output methods for cloud machine learning. Through training and training, a risk prediction model is established to obtain higher prediction efficiency. The simulation results show that the prediction accuracy of this method is 95.5%, which can estimate the hazard existing in computer information management and improve the method.

**Key words:** Machine learning algorithms, Computers, Information management system, Analytic hierarchy process

**1. Introduction.** Against the background of the rapid development in science and technology, computer information technology has introduced a rapid transformation. The reform of information management and information retrieval technologies has been greatly strengthened. In the process of office automation, the application of information technology management system has greatly improved the efficiency of office automation systems and effectively controlled the frequency of repetitive work, this can greatly improve office efficiency [16, 4, 8, 12, 7]. Computer software systems value universality and flexibility. How can we significantly improve the level of information technology and ensure the market competitiveness of our enterprise? We should analyze the development strategy of management information systems in detail. The security of computer information management systems and their business information is crucial. When there are risks to information security, some important information may be leaked. Therefore, risk prediction of computer information management systems has become a focus of current research in the field of information security. Machine learning combines knowledge of neurophysiology, computer science, and applied psychology, and there are discrepancies in analyzing large-scale data. It can be used for risk prediction in computer information management. Neffati, S. et al. The development of computer-aided design (CAD) and computer-aided manufacturing (CAM) systems in the past decade has made significant progress in biomedical and material applications. In particular, CAM and CAD systems have been applied in biomedical fields such as medical engineering, robotic surgery, medicine, and dentistry. Therefore, the accuracy and precision of CAD and CAM systems are very important for precision machining. A new brain classification CAD system is proposed based on the analysis of magnetic resonance imaging (MRI). First, the proposed reduced kernel partial least squares algorithm (DR-KPLS) is used as the feature extraction method. Then, support vector machine (SVM) is used for classification and k-fold cross validation method is used for validation. In addition, taboo search meta heuristic method is also used to determine the optimal kernel function. The algorithm named DR-KPLSVM [13]. Sun, Z, Others I believe there are still some deficiencies in the management of horse racing. The main problem is the lack of information management in the horse competition industry. Introducing information technology to the management of horse competition enterprise can achieve the management of enterprise information, make horse competition more professional and practical [17]. In order to improve the accuracy of prediction in computer information management, the author

---

\*Hebei Academy of Fine Arts, Shijiazhuang, Hebei, 050700, China (Corresponding Author).

proposes a risk prediction method based on machine learning. This method combines the Analytical Process (AHP) of machine learning and neural network to ascertain the risk of computer information management, to ensure the stability and reliability of its operation.

## 2. Methods.

**2.1. Objectives of professional teaching reform.** Since the second half of the 20th century, information industry has become the most promising new industry in the world. At present, the average growth rate of international informationization is between 15%-20%, far exceeding the growth rate of international trade [20]. In the past decade, China's information industry has developed rapidly at an average annual growth rate of 30% and has become the biggest pillar of the national economy. In the meantime, the teaching of computer major information management system must be reformed to adapt to the development of information industry, which will be of great significance. Reforming the traditional concept of vocational education, clarifying the objectives of vocational education, and positioning the training objectives of vocational education as "cultivating technical applied talents, with sufficient theoretical knowledge, focusing on skill cultivation, and quality education as the core", emphasizing characteristic education, innovative education, and lifelong education. It is necessary for students to have strong practical skills in professional and technical aspects, love their job, be diligent, and work at the forefront with peace of mind. Reforming a single talent training model, adopting a combination of academic education and certification education. The plan is to add courses related to qualification certification in the curriculum system, help students pass relevant qualification certification exams, implement a "dual certification system", and embark on a path of combining academic education and certification education. Reforming the curriculum system and updating teaching content. The new teaching plan has made two major adjustments to the curriculum system. Firstly, theoretical courses have been appropriately compressed and more practical training courses have been arranged to meet the requirements; The second is to add practical, innovative, and characteristic courses that can meet the needs of the insurance industry and employers. The outdated teaching methods and means of reform will extensively adopt modern teaching methods in professional teaching to improve classroom information capacity; By relying on laboratories and training bases, students can deepen their understanding and mastery of knowledge through experiments, and improve their skills through practical training [18, 19]. At the same time, reform the examination system, increase skill assessment, and use skill level as the main assessment indicator and evaluation criteria for students. Optimize the structure of the teaching staff, establish a "double qualified" teaching staff, reform the situation where the theoretical knowledge of the teaching staff is relatively strong and the practical knowledge is relatively weak, strengthen the training and introduction of the teaching staff, and establish a high-quality and high-level double qualified teaching staff that adapts to the characteristics of higher vocational education and has the characteristics of the times, levels, and combination of full-time and part-time, and implement goal management, project management mechanism, and incentive mechanism for teachers. Strengthen the integration of theory with practice, adopt the path of industry, academia, and research integration, reform the traditional habits of closed teaching and weak connection between theory and practice, vigorously strengthen the connection between schools and enterprises, combine learning with employment and entrepreneurship, and combine teaching, research, and industry, so that students can participate in the entire process of the real software industry to exercise during their school years.

## 2.2. Analysis of basic design principles of computer information management system.

(1) *Programming languages applied in the programming process.* The programming language commonly used by designers in computer programming is C++ [2]. However, in the actual work process of designers with high professional and technical levels, they actually do not use the C++ language. This is because the difficulty of C++ programming language is very high, so it is quite difficult for technical personnel to fully master this programming language. Moreover, it is actually quite difficult for C++ programming language to be flexibly applied, in addition, the update speed of this programming language is very slow, making it difficult to meet the actual design requirements of computer information management systems. In the actual operation of computer information management systems, it is necessary to ensure the system's running speed through the level of computer language programming. If an application wants to find highly targeted information from a database, it needs to use a data engine.

(2) *Module Design Principles.* When designing a computer information system, the most important issue involved is module design work. In the process of module design, it actually involves connecting the functional modules and hardware of the computer information management system together. Therefore, during the actual operation of the computer information management system, it can make certain guarantees for its own performance [10, 3]. Analyze this issue from a design perspective. During the process of module design work, it is generally done vertically to ensure that the system has all the functions. However, from a holistic perspective, the design of computer information management system modules is actually a horizontal design process. During the design process, the interconnectivity of each subsystem in the system should be considered, and a lot of data needs to be used, after completing the analysis and processing of data information, a detailed analysis of the actual requirements of the system can be conducted to ensure the design level. In fact, a computer information management system can be regarded as an advanced servo system, only in this way can various subsystems within the computer operate independently of each other [5]. In the process of dividing each system module, it is necessary to use a highly specialized division method. In general, all functions in a computer information system will be constructed with corresponding systems, therefore, it is possible to ensure the functionality of each system, greatly improving the operational security and stability of computer information management systems. Against the backdrop of the increasing development speed of computer information management system related technologies, the problems existing in traditional computer information management systems are gradually emerging in front of people, in order to effectively solve the problems existing in computer information management systems, the design principles of computer information management systems should be given sufficient attention, so that the probability of problem occurrence can be effectively controlled from an essential perspective.

### 2.3. Risk prediction of computer information management systems using machine learning algorithms.

(1) *Construction of Risk Prediction Index System for Computer Information Management System.* Analytic Hierarchy Process (AHP) is an information technology learning algorithm which combines the theories well and efficiently. It can monitor the initial data and make research decisions based on the analysis results. The process is very simple. Predicting the risks of the computer information management system is mainly to ensure the security of the information resources in the system. The security of information assets can be described through density preservation, integrity, reliability, authenticity, and availability [11]. Computer information management systems have openness and dynamic variability, and their risks mainly come from illegal external access to information computer information management systems, as well as illegal tampering, possession, destruction, and monitoring of information assets and application systems within the system. According to the relevant regulations on computer information management systems and information asset security protection, the risk factors of computer information management system can be divided into information assets risk, risk factor, risk factor, safety factor, and risk factor backlog. the risk factors are analyzed. Based on the five kinds of risks of computer information management, a pre-evaluation risk model for computer information management system was established using the Assessment Procedure System (AHP) (as shown in Table 2.1), which includes a total of 25 risk indicators.

(2) *Construction of risk prediction model.* Neural networks are nonlinear distributed consecutive data processing algorithms with characteristics such as self-learning preservation, transfer storage, and integration. Among them, feedforward neural networks have simple and high performance results. Among them, feedforward neural networks have good robustness and robustness. Compared with other feedforward neural networks, a hidden layer feedforward neural networks with advanced learning technology can achieve learning of neural networks under half parameter treatment, thus improving learning efficiency.  $x_p$  and  $y_p$  represents the risk indicators first and foremost risk estimates key components of the computer information management system, respectively. The data structure is described as follows:

$$J_s = \{(x_p, y_p)\}_{p-1}^s \quad (2.1)$$

In the formula,  $s$  represents the number of samples.

Table 2.1: Risk prediction index system of computer information management system

Target layer	code	Factor layer	code	Indicator layer	code			
Risk prediction of computer information management system	A	Risk factors of information assets	B1	Confidentiality	B11			
				Integrity	B12			
				Reliability	B13			
				Authenticity	B14			
				Availability	B15			
	Threatening risk factors	B2	Information is stolen, tampered with, or deleted	Network resources are destroyed	B22			
				Service suspension	B23			
				Illegal access	B24			
				Bypass control	B25			
				Authorization violation	B26			
				Vulnerability risk factors	B3	Manage Security	Hardware security	B31
							software security	B32
							Personnel safety	B33
							environmental safety	B34
							Communication security	B35
	Risk factors of safety measures	B4	Encryption measures	Anti hacker measures	B41			
				Antivirus measures	B42			
				Data backup and recovery measures	B43			
				environmental deterioration	B44			
				Service deterioration	B45			
Recovery risk factors	B5	Information recovery costs	Service recovery costs	B51				
			Information recovery costs	B52				
			Service recovery costs	B53				
			Information recovery costs	B54				
			Service recovery costs	B54				

Equation (2.1) describes the regression constraint form of extreme learning machines:

$$\min \left( \frac{1}{2} \eta_L^Y \eta_L + \frac{\gamma}{2} \phi^Y \phi \right) \quad (2.2)$$

s.t.

$$y_p = \sum_{i=1}^L \eta_i f(\delta_i x_p + z_i) - \phi_p \quad (2.3)$$

In the formula,  $L$  and  $z$  are the number of hidden layer nodes and regression error, respectively;  $f()$  and  $\phi$  are the mapping function and output layer node regression errors, respectively;  $\delta_i$  is the weight vector of  $x_p$ ,  $\eta_i$  is the weight vector representing  $y_p$ , and  $Y$  is the expected output [14, 9].

$$L(v, \phi, \eta_L) = \frac{1}{2} \eta_L^Y \eta_L + \frac{\gamma}{2} \phi^Y \phi - v(G_L \eta_L - Y - \phi) \quad (2.4)$$

Build a risk prediction model for computer information management systems:

$$y = \sum_{i=1}^L \eta_i f(\delta_i x + z_i) \quad (2.5)$$

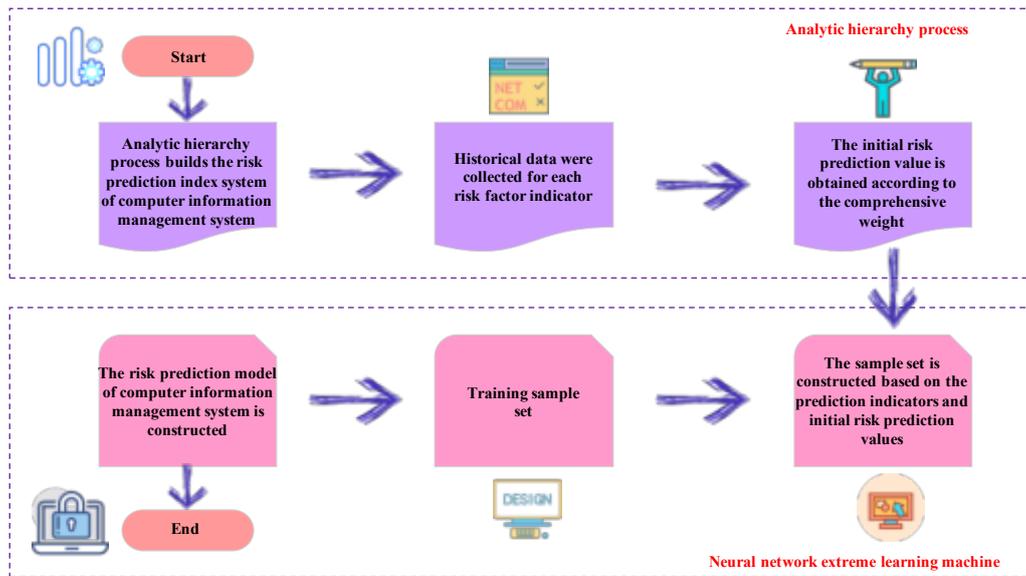


Fig. 2.1: Risk prediction workflow

The risk prediction method of the operation of computer information management system based on machine learning algorithms is illustrated in Figure 2.1. The final estimated risk value obtained by the method in Figure 2.1 can describe the risk of computer data management.

**3. Simulation Experiments.** The experiment aims to verify the effectiveness of the author’s research on risk prediction of computer information management systems. VC++600 was used to implement this study, and computer information management systems in product development departments of 20 small and medium-sized enterprises in a certain province were selected as the research objects for example verification [15].

**3.1. Risk prediction process.** In the process of predicting the risk of research items, select all research items and use the risk prediction method developed by the author, as shown in Table 2.1. After collecting the relevant data, calculated weights, and results are shown in Table 3.1.

The factors and initial risk prediction results of the selected research subjects are shown in Table 3.2. According to Tables 2.1 and 3.2, a sample set was constructed, and after training and learning, the final risk prediction result for the selected research object was 0.5878. Table 3.3 shows the comparison between the risk prediction results of all research subjects and the actual risk values studied by the authors. From Tables 3.1-3.3, it can be seen that the author’s research can effectively predict the risk of the research object [1].

**3.2. Performance verification.** By combining the author’s research in Table 3.3 with the actual risk factors, the prediction results of the author’s research results are verified. The results are shown in Figure 3.1. From Figure 3.1, it can be seen that the author studies the prediction results with high accuracy, and the prediction values are consistent with the actual risk values.

The author’s high research on predicting the risk value of research products is due to the combination of the two prediction models. It not only has the function of linear theoretical prediction models (identifying hierarchical processes), but also has the function of nonlinear theoretical prediction model (neural network cloud systems), which uses nonlinear prediction capability to fit the relationship between the evaluation indexes and the prediction results, To improve prediction accuracy [6]. By using computational information management system risk prediction, the weak links of the computer information management system can be identified based on the prediction results. Targeted solutions can be developed for different weak links to reduce system risks and ensure information security. Combine the linear method of Analytic Hierarchy Process with the nonlinear method of neural network extremum training machine, analyze the research object, fit for the relationship

Table 3.1: Weight calculation results

Factor layer	comprehensive weight	Indicator layer	comprehensive weight
5*B1	5*0.2469	B11	0.054
		B12	0.0707
		B13	0.0418
		B14	0.0582
		B15	0.0226
6*B2	6*0.2150	B21	0.0585
		B22	0.0227
		B23	0.0176
		B24	0.0528
		B25	0.0289
		B26	0.0350
6*B3	6*0.2363	B31	0.0618
		B32	0.0421
		B33	0.0464
		B34	0.0433
		B35	0.0255
		B36	0.0177
4*B4	4*0.1737	B41	0.0283
		B42	0.0348
		B43	0.0585
		B44	0.0524
4*B5	4*0.1286	B51	0.0252
		B52	0.0209
		B53	0.0395
		B54	0.0433

Table 3.2: Initial risk prediction results for each factor

Factor layer	Initial risk prediction value
B1	0.1433
B2	0.1359
B3	0.1074
B4	0.0942
B5	0.1154
Research object	0.5958

between the evaluation index and the prediction result, and make prediction risk more accurate. The changes in the two estimated estimates are shown in Table 3.4. In Table 3.4, the risk prediction results of the study subjects before and after using this method showed a significant downward trend, with a decrease in risk prediction results between 24% and 39% for each study subject. Compared with the data in the relevant statistical data of the comparative method, the risk reduction effect was more significant, indicating that the risk prediction effect of the author's study is good and can be widely promoted and used.

**4. Conclusion.** The author proposes the basic idea of educational reform, which is to design a reasonable theoretical and practical teaching system based on job knowledge requirements and centered on knowledge, ability, and quality structure. Then, a risk prediction method for computer data management based on machine learning algorithms is proposed, which combines Analytic Hierarchy Process (AHP) in machine learning algorithms with neural networks for prediction. The research results show that the author's research results

Table 3.3: Comparison of risk prediction results and actual risk values of the author’s study

Research object	Author’s research on risk prediction	Actual risk value	Research object	Author’s research on risk prediction	Actual risk value
1	0.3642	0.3827	11	0.3956	0.4569
2	0.4228	0.4802	12	0.2941	0.2762
3	0.2896	0.2685	13	0.5878	0.6032
4	0.3545	0.3757	14	0.4145	0.3856
5	0.3127	0.3127	15	0.3061	0.4094
6	0.4701	0.4523	16	0.5830	0.5813
7	0.5114	0.4997	17	0.4171	0.4027
8	0.5927	0.6024	18	0.3603	0.3603
9	0.3716	0.3860	19	0.2994	0.2995
10	0.4009	0.6594	20	0.4162	0.5369

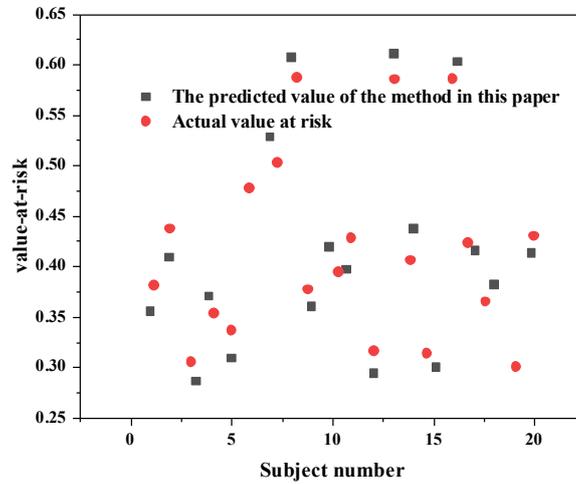


Fig. 3.1: Verification of risk prediction effect

Table 3.4: Changes in risk prediction values

number	Previous Result	Subsequent results	number	Previous Result	Subsequent results
1	0.3642	0.2241	11	0.3956	0.2500
2	0.4228	0.3146	12	0.2941	0.1949
3	0.2896	0.1929	13	0.5878	0.3956
4	0.3545	0.2243	14	0.4145	0.3002
5	0.3127	0.2362	15	0.3061	0.2107
6	0.4701	0.3248	16	0.5830	0.3863
7	0.5114	0.3776	17	0.4171	0.2912
8	0.5927	0.4209	18	0.3603	0.2233
9	0.3716	0.2503	19	0.2994	0.2205
10	0.4009	0.2895	20	0.4162	0.2936

have better prediction accuracy than the comparison method. Due to the problems in computer management systems, as well as risk issues such as personnel management and financial management, the author only focuses on risk prediction in computer information management systems. Subsequent research will mainly focus on the scalability of the author's research and expand its application fields.

## REFERENCES

- [1] N. BHANOT, J. AHUJA, H. I. KIDWAI, A. NAYAN, AND R. S. BHATTI, *A sustainable economic revival plan for post-covid-19 using machine learning approach—a case study in developing economy context*, *Benchmarking: An International Journal*, 30 (2023), pp. 1782–1805.
- [2] A. BRACA AND P. DONDIO, *Persuasive communication systems: a machine learning approach to predict the effect of linguistic styles and persuasion techniques*, *Journal of Systems and Information Technology*, 25 (2023), pp. 160–191.
- [3] Q. CHENG, S. WANG, AND X. FANG, *Intelligent design technology of automobile inspection tool based on 3d mbd model intelligent retrieval*, *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, 235 (2021), pp. 2917–2927.
- [4] X. DIAO, X. LI, S. HOU, H. LI, G. QI, AND Y. JIN, *Machine learning-based label-free sers profiling of exosomes for accurate fuzzy diagnosis of cancer and dynamic monitoring of drug therapeutic processes*, *Analytical Chemistry*, 95 (2023), pp. 7552–7559.
- [5] L. DONGMEI, *Design of english text-to-speech conversion algorithm based on machine learning*, *Journal of Intelligent & Fuzzy Systems*, 40 (2021), pp. 2433–2444.
- [6] M. HASHIMOTO AND K. NAKAMOTO, *Process planning for die and mold machining based on pattern recognition and deep learning*, *Journal of Advanced Mechanical Design, Systems, and Manufacturing*, 15 (2021), pp. JAMDSM0015–JAMDSM0015.
- [7] K. HE, Z. LI, AND D. WANG, *Overview on the design of the service and management system for field geological survey based on the remote sensing and beidou satellites*, *Journal of Geomechanics*, 18 (2022), pp. 203–212.
- [8] L. HICKMAN, C. N. HERDE, F. LIEVENS, AND L. TAY, *Automatic scoring of speeded interpersonal assessment center exercises via machine learning: Initial psychometric evidence and practical guidelines*, *International Journal of Selection and Assessment*, 31 (2023), pp. 225–239.
- [9] N. A. KUADEY, F. MAHAMA, C. ANKORA, L. BNSAH, G. T. MAALE, V. K. AGBESI, A. M. KUADEY, AND L. ADJEI, *Predicting students' continuance use of learning management system at a technical university using machine learning algorithms*, *Interactive Technology and Smart Education*, 20 (2023), pp. 209–227.
- [10] H. LI, L. ZHANG, R. GUO, H. JI, AND B. X. YU, *Information enhancement or hindrance? unveiling the impacts of user-generated photos in online reviews*, *International Journal of Contemporary Hospitality Management*, 35 (2023), pp. 2322–2351.
- [11] J. LIU, L. LIN, AND X. LIANG, *Intelligent system of english composition scoring model based on improved machine learning algorithm*, *Journal of Intelligent & Fuzzy Systems*, 40 (2021), pp. 2397–2407.
- [12] R. MA, L. ZHANG, AND Y. ZHANG, *A best possible algorithm for an online scheduling problem with position-based learning effect*, *Journal of Industrial and Management Optimization*, 18 (2022), pp. 3989–4010.
- [13] S. NEFFATI, K. BEN ABDELLAFOU, A. ALJUHANI, AND O. TAOUALI, *An enhanced cad system based on machine learning algorithm for brain mri classification*, *Journal of Intelligent & Fuzzy Systems*, 41 (2021), pp. 1845–1854.
- [14] E. PARRA VARGAS, J. PHILIP, L. A. CARRASCO-RIBELLES, I. ALICE CHICCHI GIGLIOLI, G. VALENZA, J. MARÍN-MORALES, AND M. ALCÁÑIZ RAYA, *The neurophysiological basis of leadership: a machine learning approach*, *Management Decision*, 61 (2023), pp. 1465–1484.
- [15] E. SADROSSADAT, H. BASARIR, A. KARRECH, AND M. ELCHALAKANI, *Innovative ai-based multi-objective mixture design optimisation of cpb considering properties of tailings and cement*, *International Journal of Mining, Reclamation and Environment*, 37 (2023), pp. 110–126.
- [16] Z. SARMAST, S. SHOKOUHYAR, S. H. GHANADPOUR, AND S. SHOKOOUHYAR, *Unravelling the potential of social media data analysis to improve the warranty service operation*, *Industrial Management & Data Systems*, 123 (2023), pp. 1281–1309.
- [17] Z. SUN, S. ZHANG, AND M. LIU, *Speed racing industry information management system based on machine learning and data mining*, *Journal of Intelligent & Fuzzy Systems*, 12 (2021), pp. 1–13.
- [18] Z. WANG, Z. CHENG, C. GUAN, AND H. HAN, *Intelligent information extraction algorithm of agricultural text based on machine learning method*, *Journal of Physics: Conference Series*, 1952 (2021), p. 022073.
- [19] G. XU, *Research on network intrusion detection method based on machine learning*, *Journal of Physics: Conference Series*, 1861 (2021), p. 012034.
- [20] J. YE AND W. XU, *Research on machine learning algorithm based on contour matching modal matrix*, *Journal of Physics: Conference Series*, 1883 (2021), p. 012006.

*Edited by:* B. Nagaraj M.E.

*Special issue on:* Deep Learning-Based Advanced Research Trends in Scalable Computing

*Received:* Sep 18, 2023

*Accepted:* Nov 7, 2023



## AUTOMATIC CONTROL OF LOW VOLTAGE LOAD IN POWER SYSTEMS BASED ON DEEP LEARNING

YAOHUI SUN\*, HONGYU ZHANG†, HAOLIN LI‡, SHU WANG§ AND CHUNHAI LI¶

**Abstract.** Due to the interference of false data, there is a large error in the mining results of low voltage loads in the power system. In response to this problem, the author proposes a design of an intelligent mining system for low voltage loads in the power system based on deep learning. Using ARM+DSP dual CPU structure, initializing the adapter agent, and using dual arm spiral antennas, designing a low-voltage load monitor to detect partial discharge signals in the 500-1500 MHz frequency band and suppress noise interference; By transmitting monitoring information to the intelligent switch through CAN bus or 485 bus, remote monitoring can be achieved; Based on the contact points and current characteristics of the circuit breaker, a current transformer has been designed to reduce the range of induced voltage variation; Construct a continuous set of functions MMD in the space, adjust the original network structure, establish a deep learning mining model, initial network parameters, eliminate false data in the network, optimize the network using target domain data, and combine mining engines to achieve intelligent data mining. According to the experimental results, the maximum difference between the load of phase A of the data processing system based on numerical simulation and the actual data is 1000 kVA at a time of 6 seconds; When the load of phase B is 4 seconds, the maximum difference between it and the actual data is 2000 kVA; When the load of phase C is 8 seconds, the maximum difference between it and the actual data is 2000 kVA. It has been proven that the mining error of the system is 0, and it has a precise mining effect.

**Key words:** Deep learning, Power system, Low voltage load, Intelligent mining

**1. Introduction.** The construction of the power system is an important guarantee for the well-being of the people, social stability, and national prosperity, and is an important lifeline industry. Promote green development, promote harmonious coexistence between humans and nature, accelerate the green transformation of development methods, coordinate the high-quality development of the economy and the improvement of ecological environment level, reduce energy consumption and carbon dioxide emissions per unit of GDP by 13.5% and 18% respectively [19]. This is a new requirement proposed by Premier Li Keqiang on behalf of the State Council to the energy and power industry at the Fourth Session of the 13th National People's Congress. The production and consumption of electricity are completed simultaneously, and large-scale and long-term energy storage technology is not yet mature. In power load forecasting, there is a common phenomenon of imbalance and disharmony between multiple prediction results. Without accurate and reliable load forecasting, it will result in large-scale energy waste, environmental pollution, and economic losses. In a broad sense, power load forecasting refers to the prediction of electricity consumption based on multiple influencing factors in the next hour to several years. Short term load forecasting is of great significance in ensuring the planning, reliability, and economic operation of power systems. The author mainly discusses the application of deep learning in short-term power load forecasting [16].

One of the traditional load forecasting methods is time series based forecasting methods, such as regression analysis, exponential smoothing, multiple linear regression, Autoregressive Integrated Moving Average Model (ARMA), and its improved algorithm ARIMAX. The basic idea is to predict future load values from the past and present load values of random time series. Its advantage lies in considering the temporal relationship of the data, while its disadvantage lies in the fact that the required time series is stationary and has limited predictive ability for nonlinear relational data, and has strict requirements for the stationarity of temporal data [12]. Due to the advancement of computer technology, the field of machine learning is undergoing another wave. Machine

---

\*State Grid BaiShanSupplyelectric Power Company, Baishan, JiLin, 134300, China. Corresponding author

†State Grid BaiShanSupplyelectric Power Company, Baishan, JiLin, 134300, China

‡State Grid BaiShanSupplyelectric Power Company, Baishan, JiLin, 134300, China

§State Grid BaiShanSupplyelectric Power Company, Baishan, JiLin, 134300, China

¶Shijiazhuang Kelin Electric Co., Ltd., Shijiazhuang, Hebei, 050000, China

learning algorithms are widely used in various fields such as image recognition, object detection, natural language processing, and have also achieved good results in power load forecasting. Various advanced machine learning technologies such as reinforcement learning and transfer learning have been used in load forecasting. The limitation of machine learning lies in the insufficient learning ability of high-dimensional data. Deep learning can simplify the original data of scum, extract effective information, and form features. Artificial Neural Network (ANN) is one of the fundamental algorithms of machine learning, which is a standard neural network composed of input layer, output layer, and hidden layer. Another machine learning algorithm for load forecasting is Support Vector Machine (SVM), which has good performance in small-scale data processing but poor performance in handling massive data. Deep Learning (DL) is a branch of machine learning, whose architecture is based on neural networks and more complex models. It has more hidden layers and loop structures, which endows it with stronger learning ability, adaptive ability, fault tolerance, autonomous reasoning ability, and generalization ability.

**2. References.** The Unified Power Flow Controller (UPFC) has gradually been put into current engineering applications due to its ability to quickly and independently control the active and reactive power of transmission lines, and adjust the distribution of system power flow. However, while UPFC brings many technological advantages, it also significantly increases the complexity of the power grid structure and operation mode, posing great challenges to the safe and stable operation of the power system. Therefore, studying the safety correction control of power systems containing UPFC is of great significance for ensuring the safe operation of the power grid.

At present, the conventional calculation methods for power system security correction are mainly divided into two categories, namely sensitivity based methods and optimization planning based methods. The sensitivity method generally ignores the influence of reactive power, so there is a lack of consideration for the reactive power control capability of UPFC in systems containing UPFC. In contrast, optimization planning methods can be better applied to scenarios with UPFC [4]. In summary, when considering the minimum number of adjustment components and the minimum adjustment amount as the optimization objectives for system safety correction, current research usually adopts the establishment of mixed integer nonlinear programming (MINLP) models, but these models are computationally complex and often accompanied by solveless situations. Therefore, how to improve traditional physical model based security correction methods and maximize the computational efficiency and convergence of security correction is the main problem currently facing research in this field [3].

In recent years, artificial intelligence has developed rapidly, and machine learning methods represented by deep learning have been widely favored in the field of power systems due to their ability to process large-scale data and high computational accuracy.

Song, Z. Z. J proposed a new adaptive learning deep belief network (ALDBN) with a series of growth and pruning algorithms to dynamically adjust its structure when extracting features using ALDBN. Specifically, a neuron growth algorithm considering individual and macro influences was designed to detect unstable hidden neurons, and a new hidden neuron was added around each unstable neuron to compensate for the shortcomings of local structure in feature extraction [14]. Huang, Q applies pumped storage hydropower (PSH), which can quickly track load changes, operate flexibly and reliably, balance system power, and minimize bus voltage deviation. In addition, in order to obtain the optimal control strategy for PSH, a deep reinforcement learning algorithm, namely a deep deterministic strategy gradient, is used to train the agent to solve the continuous transformation problem of the pumped storage hydroelectric wind solar (PSHWS) system [7]. Wen, T considered a multi-agent system based load frequency control method for multi-area power systems under false data injection attacks. This study can provide a better solution for load frequency control in multi region power systems under false data injection attacks. Firstly, an event triggering mechanism was introduced to determine which data should be transmitted in the controller to save limited network bandwidth. In addition, a network attack model was established using Bernoulli random variables. Then, the conditions for the system to maintain asymptotic stability under attack are given. Finally, the effectiveness of the theory proposed in this paper was verified through simulation [18].

In response to the problems existing in the traditional system mentioned above, the author proposes a design method for an intelligent mining system for low-voltage loads in power systems based on deep learning. This method can analyze the complex features hidden in low-voltage loads of high-voltage switchgear in detail,

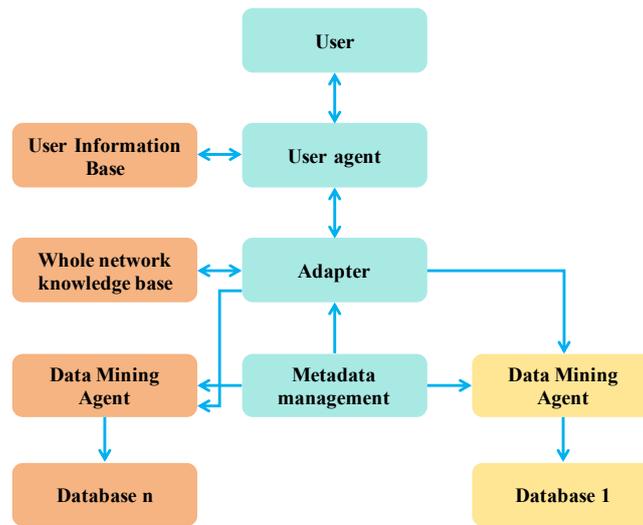


Fig. 2.1: System hardware structure

and accurately mine low-voltage loads in power systems under complex monitoring task conditions.

### 3. Methods.

**3.1. System hardware structure design.** The power system adopts advanced ARM+DSP dual CPUs, which receive three-phase current data collected by the power system monitor and send it to the station control service center through the IEC61850 protocol, achieving the remote monitoring function of the power system. The hardware structure of the system is shown in Figure 2.1.

(1) *Adapter.* The adapter is used to process the initialization information of the agent and achieve communication between the agent and the remote data collection system. This system includes the health status, location, and current system resources of each agent [11]. Each agent has an alias in the adapter, so the agent only needs to know its alias when communicating to determine where to run. Another function of this adapter is to decompose mining requests sent by UIA, and then send them separately to the corresponding DMA. After the mining process is completed, the results are directly merged and sent to UIA.

(2) *Low voltage load monitor.* The low-voltage load monitor is mainly used to monitor the insulation characteristics of the box, as shown in Figure 3.1.

From Figure 3.1, it can be seen that the online monitor is composed of a detection circuit, an IED, and an integrated information platform. It transmits partial discharge signals through high-frequency double shielded antennas, obtains as much discharge information as possible through the detection circuit, and can better suppress noise interference. Adopting a dual arm spiral high-frequency double shielded antenna, it can detect partial discharge signals in the 500 1500 MHz frequency band. Realize high-frequency dual shielding results through the IED display screen, and transmit the results to the integrated information platform through optical fiber to complete data monitoring.

(3) *Server.* The controller is designed using a dual CPU structure of ARM9 chip S3C2440A and TMS320F28335, with CAN network interface, RS-485 network interface, and RFID module set on the periphery. The IRIG-B code matching is used to provide accurate and unified time reference [2]. The high-voltage power system fully considers the impact of electromagnetic interference on equipment, and divides and isolates the power supply during the design of the power supply section. The server structure is shown in Figure 3.2.

As shown in Figure 3.2, the monitoring information is transmitted by the server to the intelligent switch through CAN bus or 485 bus. The intelligent identification unit identifies the information of the exchange device through RFID module and communicates with the service center according to the IEC61850 protocol to achieve remote monitoring function.

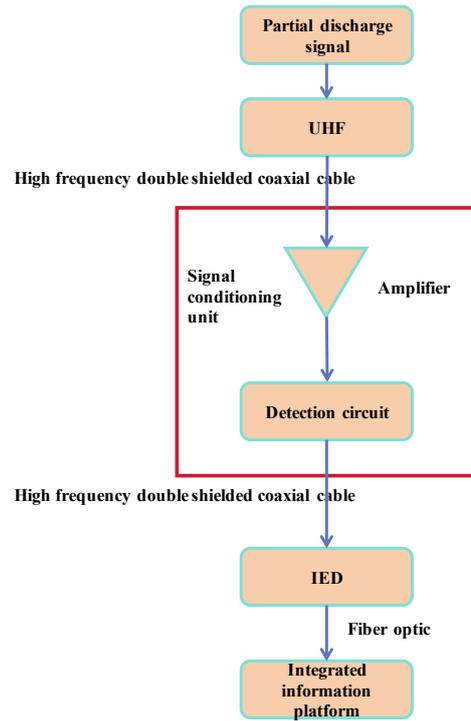


Fig. 3.1: Load data monitor

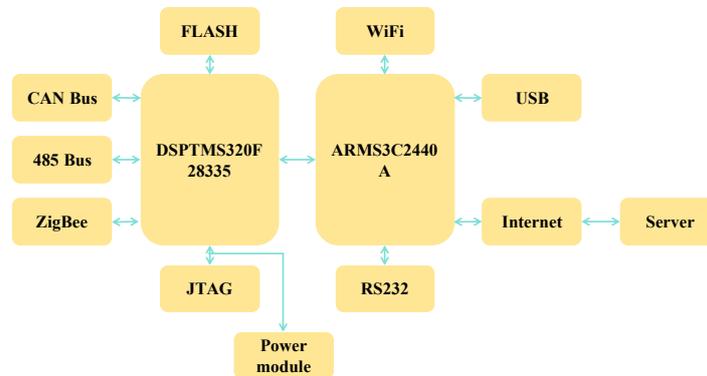


Fig. 3.2: Server structure

(4) *Power supply voltage.* Using a switching power supply module, the input AC 220 V is converted to DC 5 V, which is then converted to 3.3 V by the linear regulator chip MIC29502BU. Finally, by adjusting the chip LM1117, the 3.3 V voltage is converted to 1.5 V as the power supply. Due to the limitations of high-voltage conditions in the power system, the power supply of the monitoring device must be configured internally, using fixed energy sources such as batteries, which cannot guarantee the long-term operation of the system [5]. Therefore, a special current transformer has been designed, using permalloy as the iron core to ensure that the iron core generates high magnetic current when the excitation magnetic core is saturated at low current, thereby reducing the range of induced voltage variation, when the temperature of the busbar rises, based on the contact points and current characteristics of the circuit breaker, the principle of electromagnetic

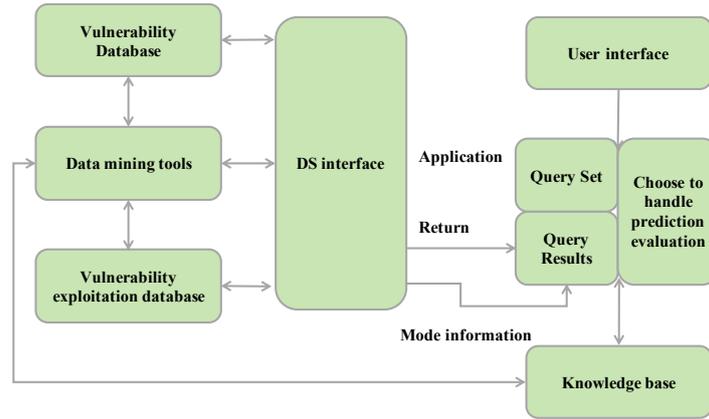


Fig. 3.3: Mining engine system structure

induction monitoring is used to directly obtain the low-voltage power supply of the equipment from the circuit breaker contact plate or busbar.

(5) *Mining Engine*. Unlike general database queries, mining engines use data mining techniques to process databases, generating specific query sets from a specific set of objects, and automatically accessing the database based on the query sets, thereby mining hidden rules in the database. The mining engine focuses on the low-voltage load library and load information library of high-voltage switchgear, achieving data mining of low-voltage loads of high-voltage switchgear. Figure 3.3 shows the system structure of the mining engine.

The mining engine takes object, domain knowledge, and pattern information as inputs, and the system generates some random query data. These query data are used as input to the database model, and the system predicts and evaluates the returned results [6].

**3.2. System software design.** Due to the differences in current distribution across different voltage and active power datasets, which affect the accuracy of mining results, it can be replaced with the weights of the fully connected layer, and the layer connection weights can also be adjusted to retain or abandon the weight feature selection of some source networks, learn new network weights to achieve the goal of retaining both source domain information and absorbing target domain information, and improve the learning ability of the network [8]. The parameters of the non adjustable layer are directly transferred and fixed from the source domain network, while the fully connected layer is retrained using target domain data, replaced and added, eliminating fixed parameters in the network to reduce learning rate, and optimizing the network using target domain data. Based on the hardware structure design of the power system, the intelligent mining process based on the above analysis is shown in Figure 3.4.

In Figure 3.4, first preprocess the source and target domain data; Then, the original deep neural network model or the trained network structure and parameters are trained using source domain data to establish a source domain classification and prediction model [15]; Finally, by analyzing the maximum mean difference between the source domain and target domain data through MMD, the distribution distance is determined, and pre training and preprocessing are carried out to complete the construction of the data mining model. Assuming that  $F$  is a continuous set of functions in the sample space, MMD can be expressed as:

$$\text{MMD}[F, p, q] = \sup_{f \in F} (E_p[f(x)] - E_q[f(y)]) \tag{3.1}$$

Assuming  $x$  and  $y$  are datasets from distributions  $p$  and  $q$ , and the size distributions of the dataset are  $m$  and  $n$ , the MMD empirical values are estimated as follows:

$$\text{MMD}[F, x, y] = \sup_{f \in F} \left( \frac{1}{m} \sum_{i=1}^m f(x_i) - \frac{1}{n} \sum_{i=1}^n f(y_i) \right) \tag{3.2}$$

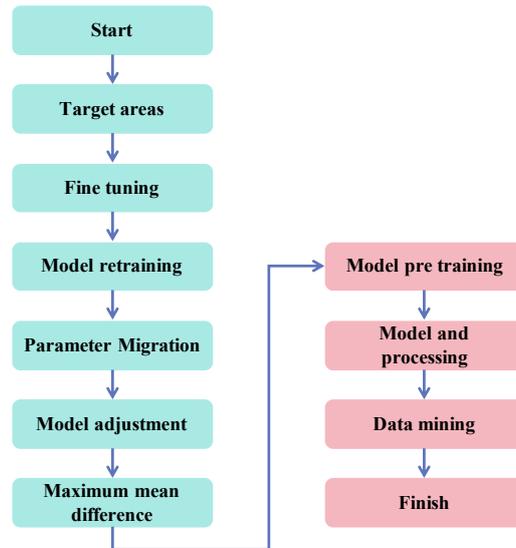


Fig. 3.4: Intelligent mining process for load data of high voltage power system

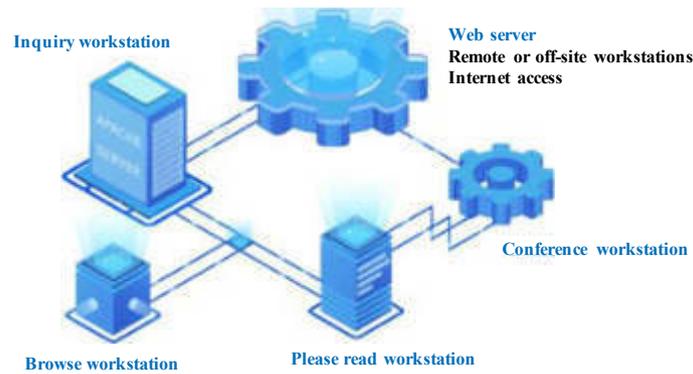


Fig. 4.1: Overall topological structure of the monitoring center

From equation (3.2), it can be seen that MMD depends on a given set of functions for a specific dataset. Under the same distribution of  $p$  and  $q$ , based on the properties of MMD, MMD is  $O$ , then it is required that  $F$  be sufficiently widespread. In order to improve the convergence speed of MMD empirical estimation, it is necessary to adjust the original network structure based on MMD to obtain a new target domain network structure, while selectively mining the parameters trained from the original network structure.

**4. Experimental Results and Analysis.** In order to verify the rationality of the design of an intelligent mining system for low voltage loads in power systems based on deep learning, experimental verification analysis was conducted.

**4.1. Experimental environment settings .** Applying the system to an airport substation in a certain city, taking the temperature of the three-phase temperature rise monitoring point in the power system as an example, the overall topology of the monitoring center is shown in Figure 4.1 for operation data monitoring from September 20, 2019 to September 24, 2019 [20].

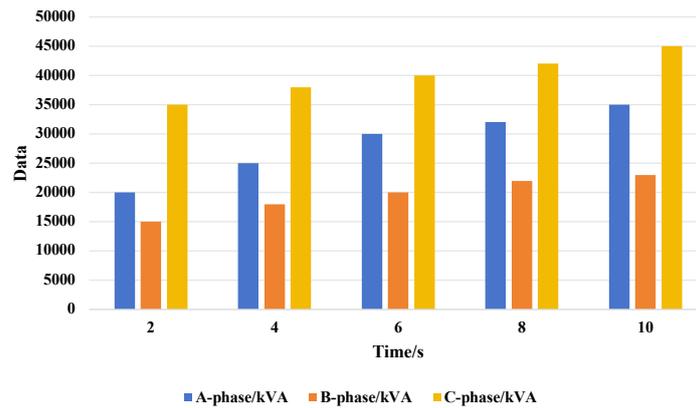


Fig. 4.2: Three phase load monitoring data

**4.2. Experimental data analysis.** Under the normal operation of various sensors in the power system, real-time three-phase load monitoring data can be effectively obtained, as shown in Figure 4.2.

**4.3. Experimental results and analysis.** Compare and analyze the mining results of three-phase low-voltage loads in power systems using machine learning based data processing systems, numerical simulation based data processing systems, and deep learning based data mining systems, as shown in Figure 4.3a-c [13].

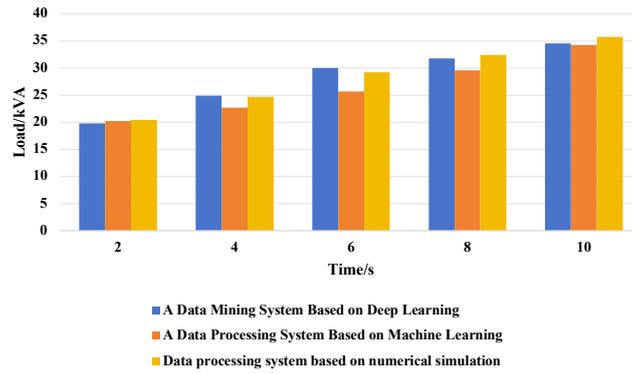
From Figure 4.3, it can be seen that the maximum difference between the load of phase A using machine learning based data processing system and the actual data is 5000 kVA at a time of 6 seconds; When the load of phase B is 10 seconds, the maximum difference between it and the actual data is 3000 kVA; The maximum difference between the load of phase C and the actual data is 3000 kVA at a time of 10 seconds [10, 1, 17, 9]. The load of phase A of the data processing system based on numerical simulation, at a time of 6 seconds, differs from the actual data by a maximum of 1000 kVA; When the load of phase B is 4 seconds, the maximum difference between it and the actual data is 2000 kVA; When the load of phase C is 8 seconds, the maximum difference between it and the actual data is 2000 kVA.

**5. Conclusion.** Verified the feasibility of applying deep learning to power system security correction. By using DNN to train and learn massive historical data, a node adjustment state recognition model with high accuracy was obtained. Greatly reducing the optimization space for node adjustment calculation, it is expected to provide good safety protection in the event of limit exceeding faults in large systems. The author has designed an intelligent mining system for low voltage loads in power systems based on deep learning. The intelligent power system not only has the functions of traditional power systems, but also has intelligent monitoring and fault diagnosis functions. It can complete corresponding operations when analyzing and processing the status of the power system locally, accurately mining low voltage loads, and laying the foundation for achieving intelligent control of power systems, however, its anti-interference function still needs to be strengthened, so future research should focus on this and strive to provide reference for relevant research in this field.

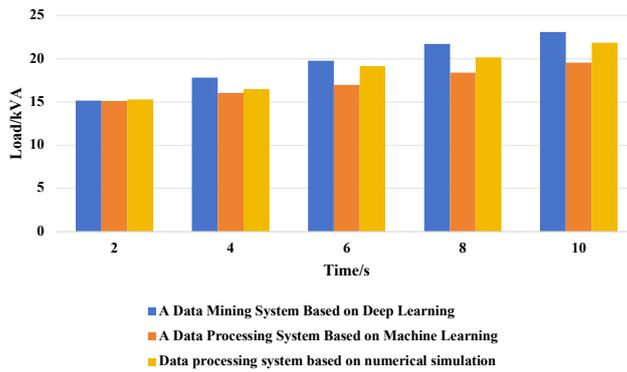
**Funding.** Science and Technology Project of State Grid Jilin Electric Power Co., LTD., 2022 (Project No.: 2022-19).

#### REFERENCES

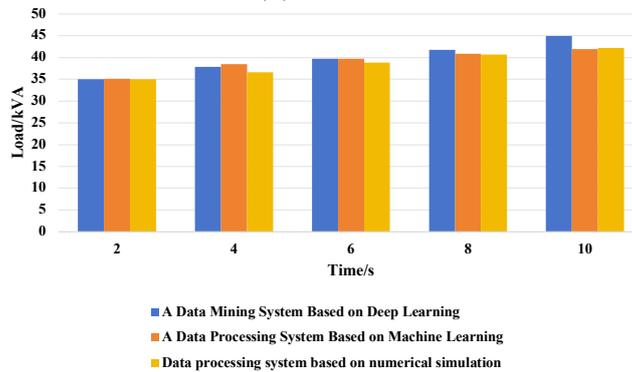
- [1] H. AHN AND H.-J. CHO, *Research of automatic recognition of car license plates based on deep learning for convergence traffic control system*, Personal and Ubiquitous Computing, 64 (2021), pp. 87–99.
- [2] W. DENG, W. PEI, N. LI, X. ZHANG, Y. YI, AND L. KONG, *Operational control of low-voltage mtdc systems in a cyber-physical environment*, CSEE Journal of Power and Energy Systems, 8 (2020), pp. 1569–1582.
- [3] M. EVERETT, Y. F. CHEN, AND J. P. HOW, *Collision avoidance in pedestrian-rich environments with deep reinforcement learning*, IEEE Access, 9 (2021), pp. 10357–10377.



(a) A-Phase



(b) B-Phase



(c) C-Phase

Fig. 4.3: Comparison of data mining results for different systems

- [4] Y. HE, X. LI, AND H. NIE, *A moving object detection and predictive control algorithm based on deep learning*, Journal of Physics: Conference Series, 2002 (2021), p. 012070.
- [5] R. M. HENRIQUES, J. A. PASSOS FILHO, AND G. N. TARANTO, *Determining voltage control areas in large scale power systems based on eigenanalysis of the qv sensitivity matrix*, IEEE Latin America Transactions, 19 (2021), pp. 182–190.
- [6] R. R. HOSSAIN, Q. HUANG, AND R. HUANG, *Graph convolutional network-based topology embedded deep reinforcement learning for voltage stability control*, IEEE Transactions on Power Systems, 36 (2021), pp. 4848–4851.
- [7] Q. HUANG, W. HU, G. ZHANG, D. CAO, Z. LIU, Q. HUANG, AND Z. CHEN, *A novel deep reinforcement learning enabled agent for pumped storage hydro-wind-solar systems voltage control*, IET Renewable Power Generation, 15 (2021), pp. 3941–3956.
- [8] R. HUANG, Y. CHEN, T. YIN, X. LI, A. LI, J. TAN, W. YU, Y. LIU, AND Q. HUANG, *Accelerated derivative-free deep*

- reinforcement learning for large-scale grid emergency voltage control*, IEEE Transactions on Power Systems, 37 (2021), pp. 14–25.
- [9] R.-H. HWANG, J.-Y. LIN, S.-Y. HSIEH, H.-Y. LIN, AND C.-L. LIN, *Adversarial patch attacks on deep-learning-based face recognition systems using generative adversarial networks*, Sensors, 23 (2023), pp. 83–90.
- [10] J. H. LEE, B. H. KIM, AND M. Y. KIM, *Machine learning-based automatic optical inspection system with multimodal optical image fusion network*, International Journal of Control, Automation and Systems, 19 (2021), pp. 3503–3510.
- [11] J. LI, S. CHEN, X. WANG, AND T. PU, *Load shedding control strategy in power grid emergency state based on deep reinforcement learning*, CSEE Journal of Power and Energy Systems, 8 (2021), pp. 1175–1182.
- [12] J. LI, T. YU, AND X. ZHANG, *Coordinated automatic generation control of interconnected power system with imitation guided exploration multi-agent deep reinforcement learning*, International Journal of Electrical Power & Energy Systems, 136 (2022), p. 107471.
- [13] ———, *Coordinated load frequency control of multi-area integrated energy system using multi-agent deep reinforcement learning*, Applied Energy, 26 (2022), pp. 1865–1885.
- [14] W. SONG, S. ZHANG, Z. WEN, AND J. ZHOU, *A novel adaptive learning deep belief network based on automatic growing and pruning algorithms*, Applied Soft Computing, 104 (2021), pp. 69–102.
- [15] X. SUN AND J. QIU, *A customized voltage control strategy for electric vehicles in distribution networks with reinforcement learning method*, IEEE Transactions on Industrial Informatics, 17 (2021), pp. 6852–6863.
- [16] H. TANG, S. WANG, K. CHANG, AND J. GUAN, *Intra-day dynamic optimal dispatch for power system based on deep q-learning*, IEEE Transactions on Electrical and Electronic Engineering, 16 (2021), pp. 954–964.
- [17] W. TANG, Q. YANG, X. HU, AND W. YAN, *Deep learning-based linear defects detection system for large-scale photovoltaic plants based on an edge-cloud computing infrastructure*, Solar Energy, 231 (2022), pp. 527–535.
- [18] T. WENG, Y. XIE, G. CHEN, Q. HAN, Y. TIAN, L. FENG, AND Y. PEI, *Load frequency control under false data inject attacks based on multi-agent system method in multi-area power systems*, International Journal of Distributed Sensor Networks, 18 (2022), pp. 4610–4618.
- [19] L. YIN, C. ZHANG, Y. WANG, F. GAO, J. YU, AND L. CHENG, *Emotional deep learning programming controller for automatic voltage control of power systems*, IEEE Access, 9 (2021), pp. 31880–31891.
- [20] L. ZHU AND Y. LUO, *Deep feedback learning based predictive control for power system undervoltage load shedding*, IEEE Transactions on Power Systems, 36 (2021), pp. 3349–3361.

*Edited by:* B. Nagaraj M.E.

*Special issue on:* Deep Learning-Based Advanced Research Trends in Scalable Computing

*Received:* Sep 18, 2023

*Accepted:* Nov 11, 2023



## TARGET IMAGE PROCESSING BASED ON SUPER-RESOLUTION RECONSTRUCTION AND DEEP MACHINE LEARNING ALGORITHM

YANG LIN<sup>\*</sup>, PING ZHANG<sup>†</sup>, HE ZHANG<sup>‡</sup> AND GUOPING SONG<sup>§</sup>

**Abstract.** In dictionary-based single-frame image reconstruction algorithms, dictionaries rely on the design of artificial shallow features and are limited in their ability to represent image features. Therefore, this paper proposes a high-accuracy reconstruction method based on deep learning feature dictionary. This algorithm first uses a deep network to learn high-resolution and low-resolution training example images with deep features; Then co-train the feature dictionary under the super dense framework of the sparse dictionary; Finally, a single low-resolution image can be input and a super-resolution reconstruction can be performed using a dictionary. From the theoretical analysis, the introduction of deep network to extract the deep-level features of the image and its use in dictionary training is more beneficial to complement the high-frequency information in the low-resolution image. Experiments show that the proposed method achieves the best results in terms of both the peak signal-to-noise ratio and the gradient energy function of the reconstructed images. This shows that compared with traditional interpolation methods and some deep learning methods, the proposed method can recover image details to a high degree while preserving the original image damage information. This proves that the subjective visual and objective evaluation indicators of the algorithm presented in this article are higher than those of the comparative algorithm.

**Key words:** Deep learning, Dictionary learning, Super resolution, Deep level feature extraction, Single frame image

**1. Introduction.** In the era of the Internet, information transmission is essential. The transmission methods include sound, video, images, text, etc. The importance of images in information transmission is self-evident. When computers process information, the source of images is generally composed of photographic equipment based on spatial sampling and quantized values [8]. The amount and credibility of information displayed in an image are closely related to its resolution. Image size refers to the number of pixels in an image, while resolution refers to the pixel density per unit. At the same time, the semantic information expression is more complete. The above information is a key element in a series of operations such as image analysis, processing, and classification. In the field of computer vision, image analysis and processing occupy an important position. Image classification, facial recognition, autonomous driving, etc all belong to the category of image processing, with the difference being that single frame and multi frame images have different processing speed requirements. But when doing image processing, a clear image is very important, and the image super-resolution here can be used as image preprocessing and other work to reconstruct clarity.

Image super-resolution reconstruction is the process of processing low resolution images through certain algorithms to obtain high-resolution or ideal resolution images. High resolution images, due to their higher pixel density, can display more detailed information and present finer image quality. Using more powerful photography equipment can help obtain higher resolution images, but as a result, costs cannot be effectively controlled. So the super-resolution algorithm plays a huge role in saving costs and obtaining high-resolution images [10].

At the same time, using high-resolution images for network transmission can cause problems such as slow transmission speed, network congestion, and even transmission failure. In this way, both the smoothness of task completion and user satisfaction cannot be solved. So we can use smaller low resolution images during transmission and restore them using super-resolution algorithms on the client side, which can greatly alleviate network pressure and compression costs. In addition, many industries use super-resolution algorithms for end-to-end image processing, and industries that require image information such as satellite and drone photography,

---

<sup>\*</sup>Jilin Distance Education Technology Innovation Center, Changchun, Jilin, 130022, China (Corresponding Author)

<sup>†</sup>Jilin Distance Education Technology Innovation Center, Changchun, Jilin, 130022, China

<sup>‡</sup>Jilin Distance Education Technology Innovation Center, Changchun, Jilin, 130022, China

<sup>§</sup>Jilin Distance Education Technology Innovation Center, Changchun, Jilin, 130022, China

driving information collection, and medical examinations have high popularity of super-resolution algorithms [19]. For example, it is necessary to magnify confidential photos taken by spy satellites, distinguish distant object categories from front and rear car cameras, capture clear license plate numbers from intersection cameras, and capture details of distant scenery by photographers, all of which require image super-resolution technology. In addition, it can also help us deal with a series of problems such as noise interference and motion blur when using images collected through low-end equipment. From this, it can be seen that image super-resolution reconstruction has a wide range of applications and high value. While it brings help to various industries, it also has improvement and research value [1].

**2. Reference.** Super-resolution imaging technology, which aims to create a high-resolution image from a single low-resolution image or multiple low-resolution images, has become a useful tool for improving the accuracy of nuclear resonance imaging. Traditional image solvers often use translation methods such as bilinear translation, bicubic translation, and nearest neighbor translation. However, because the adjacent pixel values in the image are not constant, the image returned by the definition process is blurred [12]. Ultra-high-resolution multi-images can extract information from low-resolution multi-images to create high-resolution images, but the reconstruction process requires complex registration work, high time and space, and is not easy, which is not seen in practical applications.

Deep learning is one of the many learning machines that have made significant progress in ultra-high-resolution image processing, playing an important role in tasks such as object tracking, satellite imagery, and medical imaging. Deep learning involves training multiple image models, building a learning neural network, extracting correlations between high-resolution images and low-resolution images, and using this correlation graph to obtain the frequency information needed for updated images [14].

The main purpose of super-resolution images is to recover high-quality or high-resolution images from bad samples of low-quality or low-resolution images. In this research project, Muhammad, W proposed a new concept to reduce the number of redundant networks and improve operational speed, inspired by ResNet and Xception networks [13]. Fukami K et al. proposed a new method for data generation inspired by super-resolution and intermediate techniques using machine learning techniques to control recovery through spatio-temporal data flow. For the current machine learning based on updated data, we use neural network-based descending cross-connections/multiple models to integrate water mixture properties into their network structure [2]. The design of remote sensing surveys requires the use of special satellite images. However, these images may not be available at any time. Therefore, the modification method requires remote viewing of the image at that time. In this paper, Isa, SM proposed a machine learning model to transform Landsat-8 images into Sentinel-2 images. The inspiration for this model comes from the development of super-problem models based on deep learning. Studies have shown that the proposed model can predict Sentinel-2 images that are quantitatively and qualitatively similar to the original images [6].

In addition, multiple research groups have used different deep learning neural networks to perform three-dimensional virtual reconstruction of fluorescence microscope images to accurately locate the axial position of single molecules, or to better explore various life processes of cells through color separation processing of polychromatic images.

### 3. Methods.

#### 3.1. Related theories.

**3.1.1. Principle and limitations of ScSR algorithm.** At present, the Sparsecoding superresolution (ScSR) algorithm is the most classic among dictionary building and learning based superresolution algorithms. This algorithm uses a natural image library as the training dataset to train dictionary pairs  $D_h$  (High Resolution Dictionary) and  $D_l$  (Low Resolution Dictionary) that match high-resolution (HR) and low-resolution (LR) images; Then, for each input low resolution block  $y$ , find its sparse representation coefficient  $\tilde{a}$  in  $D_l$ , and the corresponding high resolution feature blocks based on  $D_h$ . will be combined through these coefficients [9]; Finally, output high-resolution feature block  $x$ . The reconstruction principle of this algorithm can be expressed as

$$\tilde{a} = \min_a \|\tilde{D}a - \tilde{y}\|_2^2 + \lambda \|a\|_1 \quad (3.1)$$

In the formula:

$$\tilde{D} = \begin{bmatrix} FD_1 \\ \beta PD_h \end{bmatrix}, \tilde{y} = \begin{bmatrix} Fy \\ \beta\omega \end{bmatrix},$$

$D$  represents an overcomplete dictionary,  $F$  represents a feature extraction operator. Using coefficient  $\tilde{a}$  and dictionary  $D_h$ , the HR image block is obtained, which can be represented as

$$x = D_h \hat{a} \quad (3.2)$$

By combining the HR image blocks of various buildings, the HR image can be calculated. However, the accuracy of the predicted HR image is related to the accuracy of the dictionary atoms involved in the superresolution, and the key is whether the features of the extracted image describe the image pattern. In the ScSR algorithm, high-sensitivity features sensitive to the human eye are selected for image feature extraction and representation, and high-sensitivity features are extracted in the first step and the second step. The gradient operator is

$$\begin{cases} f_1 = [-1, 0, 1] \\ f_2 = f_1' \\ f_3 = [-1, 0, -2, 0, 1] \\ f_4 = f_3' \end{cases} \quad (3.3)$$

In the formula,  $f_1$  and  $f_3$  extract the horizontal gradient features  $f_2$  and  $f_4$  of the image, and extract the vertical gradient features of the image. The four gradient features are ordered as a vector based on the block features of the LR image. During the reconstruction of the over-resolved problem, the ScSR algorithm selects large gradient feature values as candidate dictionary atoms [4].

However, for some texture images with complex data sets, the change in texture gradient value is smaller than the edge pattern gradient value. If the atoms of the candidate dictionary are selected according to the above properties, the resulting dictionary will be focused on the edge structure and thus the structure of the building image.new and rare super-resolution images [5]. In addition, when there is a difference in image type between the super-resolution reconstruction image and the training set image, the candidate dictionary atoms are selected according to structure or the edge gradient value during super-resolution reconstruction is inappropriate phase, thus affecting the image quality of the super-resolution reconstruction. Therefore, the ScSR supersolving algorithm, which constructs a dictionary according to pseudo-rules, is based on a shallow dictionary, and the dictionary has some global constraints. Deep training of image models is a good way to improve the quality of reconstructed images to create a deep dictionary and improve image feature representation.

**3.1.2. Deep learning feature dictionary.** Features are an important factor in machine learning algorithms and have a decisive impact on the final model. If the training sample images have good feature expression ability, usually the computational model can obtain satisfactory results. In recent years, with the increasing complexity of target images, people have begun to pay attention to how to construct more efficient feature dictionaries to achieve more accurate descriptions of image edge information, rich textures, and geometric structures [7]. Deep networks can mine deep knowledge of data, extract deeper features of training sample images using deep networks, break through the limitations of artificial rule features, and improve the expression ability and adaptability of dictionaries.

Therefore, this paper combines the advantages of PCANet deep network and ScSR distributed dictionary algorithms to propose a high-resolution image restoration algorithm based on PCANet dictionaries. The image blocking algorithm is shown in Figure 3.1. In the training phase, it is estimated that the features of low-resolution and high-resolution images have similar expressions in their book translation. Then, a PCANet deep network is used to obtain the deep level features of the training sample images, and a pair of overcomplete feature word dictionaries  $D_h$  and  $D_l$  are obtained through dictionary joint training. Among them,  $D_h$  represents a high-resolution image feature dictionary, and  $D_l$  represents a low-resolution image feature dictionary [20].

In the super-resolution reconstruction stage, the PCANet method is also used to extract deep level features from low resolution images, solve for the sparse representation of each low resolution feature block in the

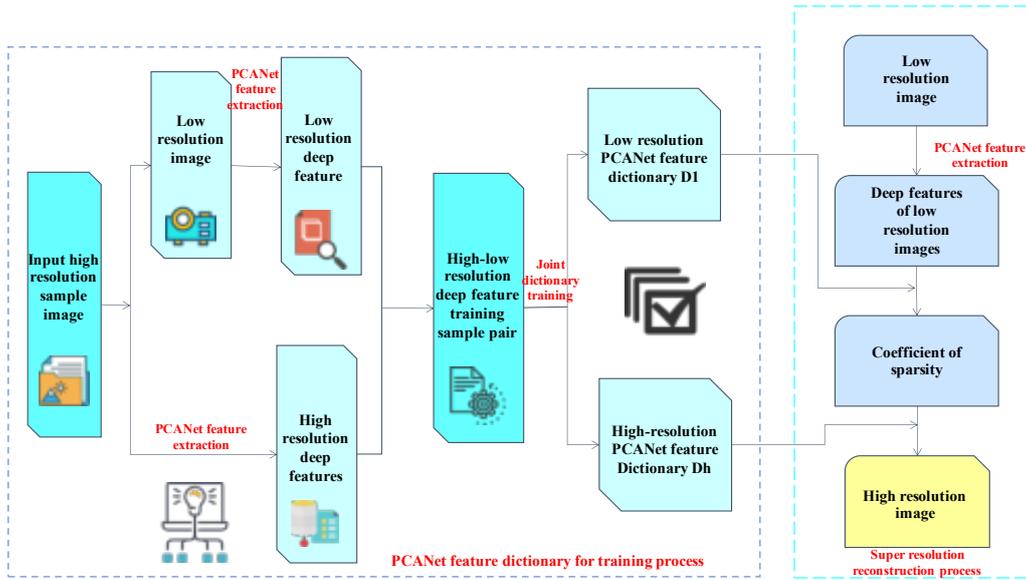


Fig. 3.1: Overview of single image superresolution algorithm based on PCANet feature dictionary

dictionary  $D_l$  of the image to be super-resolved, and directly apply the sparse representation coefficient of the low resolution image feature block on  $D_l$  to  $D_h$ , thereby obtaining the corresponding high-resolution image feature blocks, and ultimately achieving super-resolution reconstruction of the low resolution image. This algorithm uses PCANet deep network mining to train sample image features, which can obtain richer feature information than non deep networks. The deep level feature dictionary constructed on this basis also improves the expression ability of the feature dictionary, achieving effective improvement in the quality of reconstructed images [17].

### 3.2. Algorithm in this article.

**3.2.1. Image preprocessing.** Firstly, blur down sample  $K$  high-resolution images in the training data, and then enlarge them to the same size as the high-resolution image to obtain a low resolution image corresponding to the high-resolution image, forming a training sample pair:  $T = \{X_h, X_1\}$ , where  $X_h = \{x_h^i\}_{i=1}^K$  is the high-resolution feature and  $X_1 = \{x_1^i\}_{i=1}^K$  is the low resolution feature. Then, calculate the blocking matrix for each training sample. Taking high-resolution training samples as an example, select a  $k_1 \times k_2$  sliding window (usually a square window with a side length of 3, 5 or 7 pixels), and each  $E * E$  sized image will be transformed into  $m \times n \quad k_1 \times k_2$  sized image blocks after extracting local features through the sliding window; Then average these  $m \times n$  image blocks and complete the feature extraction operation for a single image. Perform the above operation on all  $N$  high-resolution images to obtain a new data matrix  $X$  with  $N \times m \times n$  columns. Each column of the matrix represents an image block with a total of  $k_1 \times k_2$  elements [11].

The  $i$ -th high-resolution training sample obtained can be expressed as equation (3.4), the block matrix of the overall high-resolution sample can be expressed as equation (3.5), and the block matrix calculation of the low resolution sample is the same as that of the high-resolution sample, expressed as equation (3.6).

$$\bar{X}_{hi} = [\bar{x}_{hi,1}, \bar{x}_{hi,2}, \dots, \bar{x}_{hi,mn}] \tag{3.4}$$

$$X_h = [\bar{X}_{h1}, \bar{X}_{h2}, \dots, \bar{X}_{hK}] \in R^{k_1 k_2 \times Kmn} \tag{3.5}$$

$$X_1 = [\bar{X}_{11}, \bar{X}_{12}, \dots, \bar{X}_{1K}] \in R^{k_1 k_2 \times Kmn} \tag{3.6}$$

**3.2.2. Feature extraction.** Based on the above training sample matrix, PCANet feature extraction is performed, and this feature is used as the training sample feature in the ScSR model for PCANet feature

dictionary training. The PCANet deep learning network completes image feature extraction in three layers. Taking high-resolution training images as an example, a detailed feature extraction process is provided below.

(1) *1st layer.* The first layer feature extraction process is the process of constructing a PCA filter and performing convolution, as shown in the first rectangular box (Firststage) in Figure 3.1. Assuming that the number of filters required in the  $i$ -th layer is  $L_1$ , a series of standard orthogonal matrices are searched for to minimize the reconstruction error through equation (3.7). By extracting the eigenvectors corresponding to the first  $L_1$  maximum eigenvalues of the covariance matrix  $X_h$ , a feature mapping matrix is formed, which is the PCA filter, as shown in equation (3.8).

$$\min_{V \in R^{k_1 k_2 \times L_1}} \|X_h - VV^T X_h\|_F^2 \quad S.T. V^T V = I_{L_1} \quad (3.7)$$

$$W_{hL}^1 = \text{mat}_{k_1 k_2} (q_{hL} X_h X_h^T) \in R^{k_1 k_2} \quad L = 1, 2, \dots, L_1 \quad (3.8)$$

Rearrange each column of the  $L_1$  feature vectors to obtain 1 patch, which results in an  $L_1$  of  $k_1 \times k_2$  window. Then, for each image, perform a convolution using these  $L_1$  windows [18]. The main information of the training sample image can be retained through the first layer PCA filter, represented as

$$I_{hi}^L = I_{hi}^L \times W_{hL}^1 \quad i = 1, 2, \dots, K \quad (3.9)$$

(2) *2nd layer.* The PCA mapping process of the second layer is shown in the second rectangular box (Secondstage) in Figure 3.1. Similar to the blocking operation of the sample image in the first layer, the output result of the PCA mapping in the first layer is used as the input of the second layer. Similarly, the matrix is subjected to block sampling, cascading, and zero averaging operations in the second layer, and the results are represented as follows

$$Y_h^L = [\bar{Y}_{h1}^L, \bar{Y}_{h2}^L, \dots, \bar{Y}_{hK}^L] \in R^{k_1 k_2 \times K_{\bar{m}\bar{n}}} \quad (3.10)$$

$$Y_h = [Y_h^1, Y_h^2, \dots, Y_h^{L_1}] \in R^{k_1 k_2 \times L_1_{\bar{m}\bar{n}}} \quad (3.11)$$

Similarly, the PCA filter consists of eigenvectors corresponding to the covariance matrix, and the filter is

$$W_{hL}^2 = \text{mat}_{k_1 k_2} (q_{hL} (Y_h Y_h^T)) \in R^{k_1 k_2} \quad L' = 1, 2, \dots, L_2 \quad (3.12)$$

Due to the presence of  $L_1$  filter kernels in the first layer, the second layer performs the same step of feature extraction for each feature output from the previous layer, resulting in  $L_1$  feature outputs. Finally, for each sample image, PCANet will output  $L_1 \times L_2$  feature matrices, as shown in equation (3.13). In terms of structure, the feature extraction process of the two-layer PCA is very similar, and PCANet can also be expanded into a deep network structure containing more layers as needed [3].

$$O_{hi}^{L'} = \{I_{hi}^L \times W_{hL'}^2\}_{L'=1}^{L_2} \quad (3.13)$$

(3) *Output layer.* The output layer mainly targets each output matrix in the second layer, which is binarized to only contain integers 1 and 0. Then, the matrix is binarized and hashed, as shown in the formula:

$$T_{hi}^L = \sum_{L'=1}^{L_2} 2^{L'-1} H(O_{hi}^{L'}) = \sum_{L'=1}^{L_2} 2^{L'-1} H(I_{hi}^L \times W_{hL'}^2) \quad L = 1, 2, \dots, L_1 \quad (3.14)$$

In the formula,  $T_{hi}^L$  is the result of hash encoding of high-resolution training image features,  $2^{L'-1}$  is the transformation coefficient that causes each pixel value in the image to become a numerical value between 0 and 255. The function  $H(\cdot)$  is similar to a unit step function, which quantifies and increases the difference between each feature. Then, the above results are histogram encoded to complete PCANet feature extraction of a high-resolution sample image, and the extracted results can be represented as

$$F_{hi} = [\text{Bhist}(T_{hi}^1), \dots, \text{Bhist}(T_{hi}^{L_1})]^T \in R^{(2^{L_1})_{L_1 B}} \quad (3.15)$$

Similarly, it can be inferred that after the same steps as high-resolution training images, the PCANet feature  $F_{li}$  of low-resolution training images can be represented as

$$F_{li} = \left[ \text{Bhist} (T_{li}^1), \dots, \text{Bhist} (T_{li}^{L_1}) \right]^T \in R^{(2^{L_1})L_1 B} \quad (3.16)$$

In the formula,  $F_{hi}$  and  $F_{li}$  represent the feature extraction results of the sample image; Bhist represents the histogram encoding process;  $B$  represents the number of segmented image blocks in the sample image [16].

**3.2.3. Dictionary training.** This article adopts a joint sparse encoding approach to train dictionaries under the ScSR framework. The goal of the algorithm is to obtain a pair of dictionary pairs  $D_h$  and  $D_l$  that can describe the complex features of the sample, so that the PSAllet features  $F_{hi}$  and  $F_{li}$  generated by  $K$  on the image have the same sparse representation on  $D_h$  and  $D_l$ , and  $F_{hi}$  and  $F_{li}$  have the same representation coefficients, i.e

$$\{D_h, a\} = \underset{D_h, a}{\operatorname{argmin}} \|F_{hi} - D_h \cdot a\|_2^2 + \sum_{i=1}^K \lambda_i \|a_i\|_1 \quad (3.17)$$

$$\{D_l, a\} = \underset{D_l, a}{\operatorname{argmin}} \|F_{li} - D_l \cdot a\|_2^2 + \sum_{i=1}^K \lambda_i \|a_i\|_1 \quad (3.18)$$

In order to ensure that high-resolution image features and low-resolution image features have the same sparse expression form regarding their respective dictionaries, a joint training strategy is adopted for equations (3.17) and (3.18), namely

$$\{D_l, D_h, a\} \underset{D_l, D_h, a}{\operatorname{argmin}} \|F_{hi} - D_h \cdot a\|_2^2 + \frac{1}{M} \|F_{li} - D_l \cdot a\|_2^2 + \left( \frac{1}{N} + \frac{1}{M} \right) \sum_{i=1}^K \lambda_i \|a_i\|_1 \quad (3.19)$$

In the equation,  $N$  and  $M$  represent the dimensionality of rearranging the element values of high and low resolution feature image blocks into column vectors, respectively.  $1/N$  and  $1/M$  are used to balance the costs of  $D_h$  and  $D_l$  in equations (3.17) and (3.18). For convenience in solving, equation (3.19) is written as

$$\{D_c, a\} = \underset{D_c, a}{\operatorname{argmin}} \|F_c - D_c \cdot a\|_2^2 + \sum_{i=1}^K \lambda'_i \|a_i\|_1 \quad (3.20)$$

$$D_c = \begin{bmatrix} \frac{1}{\sqrt{N}} D_h \\ \frac{1}{\sqrt{M}} D_l \end{bmatrix}, X_c = \begin{bmatrix} \frac{1}{\sqrt{N}} F_h \\ \frac{1}{\sqrt{M}} F_l \end{bmatrix}, \lambda'_i = \left( \frac{1}{N} + \frac{1}{M} \right) \lambda_i \quad (3.21)$$

Equation (3.21) is solved using an iterative method. Firstly, by giving a dictionary  $D_c$ , solve for each pair of training sample data  $F_{ci}$ ; The sparse representation coefficient  $a_i$  on  $D_c$  is obtained, resulting in a sparse representation matrix  $\alpha = \{a_i\}_{i=1}^K$ . Finally, the dictionary  $D_c$  is updated based on  $\alpha$ .

**3.2.4. Image reconstruction.** After obtaining the high resolution and low resolution image feature dictionaries for  $D_h$  and  $D_l$ , it is necessary to solve the classical optimization problem for each low resolution test image

$$\hat{a} = \min_a \|\tilde{D}a - \tilde{y}\|_2^2 + \lambda \|a\|_1 \quad (3.22)$$

In the formula:

$$\tilde{D} = \begin{bmatrix} FD_1 \\ \beta PD_h \end{bmatrix}, \tilde{y} = \begin{bmatrix} Fy \\ \beta \omega \end{bmatrix},$$

$F$  is the PCANet feature,  $P$  is the overlap area between the current high-resolution image block and the reconstructed high-resolution image block. This strategy can reduce blocky effects,  $\omega$  is the number of pixels

in the overlap area, and  $\beta$  measures the matching between the input low-resolution feature block and the high-resolution overlap area. Solve equation (3.22) to obtain the coefficient expression coefficient  $\alpha$  for each low resolution image, and apply it to  $D_h$  to obtain the initial high-resolution feature information image  $X_0 = D_h\alpha$ . Solve the following optimization problem for  $X_0$  to ensure that the final high-resolution image meets the reconstruction constraints, namely

$$X^* = \min_X \|HX - Y\|_2^2 + \lambda \|X - X_0\|_2^2 \quad (3.23)$$

In the equation,  $H$  is the image degradation operator ( $H$  is related to the imaging process, and in this experiment, this operator only represents downsampling of the test image). By using the gradient descent method to solve equation (3.23), a super-resolution result image can be obtained.

**3.2.5. Algorithm steps.** The algorithm in this article first extracts PCAdet features from low resolution images, and trains dictionaries based on these deep level features to achieve super-resolution reconstruction. The algorithm is shown below.

---

**Algorithm 1** Single frame image super-resolution reconstruction based on PCAdet feature dictionary

---

**Input:** Low resolution image  $Y$ , high and low resolution dictionary pairs  $D_h, D_l$

**Output:** High-resolution image  $X^*$

- 1: Perform  $2x$  unsampling interpolation on low resolution images with  $Y' = S(Y)$ , where  $S$  is the unsampling operator
  - 2: Extracting PCANet Features  $F_1$  from Low Resolution Images
  - 3: Establish the relationship between dictionary  $D_h, D_l$ , and  $F_1$  according to equation (3.22)
  - 4: **for**  $i = 1$  to  $T$  **do**
  - 5:     Calculate  $\hat{a} = \min_a \|\tilde{D}1a - \tilde{y}\|_2^2 + \lambda \|a\|_1$  to obtain the sparsity coefficient  $\hat{a}$  of the low resolution PCANet feature with respect to the low resolution dictionary  $D_l$
  - 6:     Based on the sparsity coefficient  $\hat{a}$  and high-resolution dictionary  $D_h$ , estimate the feature information of high-resolution images as  $X_0 = D_h\alpha$
  - 7:     Synthesize the obtained high-resolution image feature information with the low-resolution image to obtain the initial super-resolution result image
  - 8: **end for**
  - 9: Using the gradient descent method, calculate  $X^* = \min_X \|HX - Y\|_2^2 + \lambda \|X - X_0\|_2^2$ . Obtain the super-resolution result image  $X^*$  that is closest to the initial high-resolution image
- 

## 4. Experimental Results and Analysis.

**4.1. Experimental dataset.** At present, there are few publicly available breast cancer MRI image data sets, so this experimental group has built a breast cancer MRI image data set. This dataset is provided by Liaoning Cancer Hospital and includes a total of 260 cases, including 17 male cases and 243 female cases, covering the age range of 32-78 years. In the experiment, CE3 images of patients were used, with a total of 48 slices per case. All black and poor quality slice images were removed from the slices, resulting in a self built dataset of 1310 slices. The obtained image is downsampled using the bicubic downsampling method at a rate of 4 times, resulting in a low resolution image.

**4.2. Training details and parameters.** Before training, the images were standardized to accelerate the convergence speed of the network. During training, a low resolution image with a size of  $128 \times 128 \times 3$  was input, and the number of training sessions per batch was set to 4. The optimizer selected Adam and set the learning rate to  $1 \times 10^{-4}$ . The loss function was selected as Mean Absolute Error (MAE). The training process was conducted 1200 iterations, lasting approximately 34 hours. All experiments were conducted using the Keras framework on NVIDIA RTX2080ti [15].

Table 4.1: Image subjective evaluation form

Grade	Absolute measurement scale	Detail	Score
1	Excellent	The best in the group	10
2	Good	Better than the average in the group	8
3	Average	Group average	6
4	Fair	Worse than the average in the group	4
5	Poor	Worse in the group	2

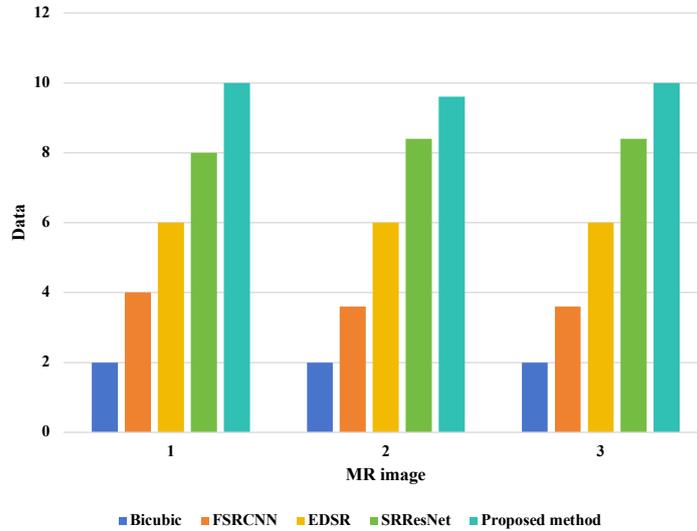


Fig. 4.1: Comparison of subjective evaluation values of super-resolution reconstruction methods

**4.3. Experimental results and discussion.** After the training is completed, high-resolution images are downsampled 4 times to obtain low resolution images as test images. The low resolution images are then input into FSRCNN, EDSR, SRResNet, and the proposed network for testing. In addition to deep learning methods, the images are also subjected to bicubic interpolation. In order to qualitatively analyze the quality of image super-resolution, image quality was evaluated from both subjective and objective perspectives. Subjectively, the reconstructed image quality under different methods is evaluated based on the subjective evaluation table of the image. This evaluation method divides the image into 5 levels based on its relative and absolute quality, and the relative and absolute measurement scales of each level are shown in Table 4.1. According to this evaluation indicator, the volunteers rated the reconstructed images, and the results are shown in Figure 4.1. From Figure 4.1, it can be seen that compared to the images recovered by other methods, the reconstructed images by the proposed method are more in line with human subjective visual perception, and the restoration effect of image details and textures is better than other methods.

Objectively, PSNR and gradient energy values were selected as the main evaluation indicators. Figures 4.2 and 4.3, respectively represent the PSNR and gradient energy values of each image under different reconstruction methods. From Figures 4.2 and 4.3, it can be seen that the proposed model achieves the best results in peak-to-noise ratio and gradient power function of the reconstructed image. It shows that compared with traditional annotation methods and some deep learning methods, the proposed model can recover the maximum amount of image details while preserving the context information.

Figure 4.4a-c shows the comparison of the super-resolution reconstruction results PSNR, SSIM, and computational cost (Time) for each experiment. From the PSNR numerical results, the interpolation method has

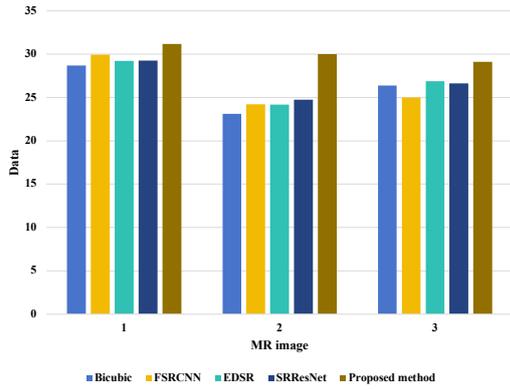


Fig. 4.2: Comparison of PSNR values of various super-resolution reconstruction

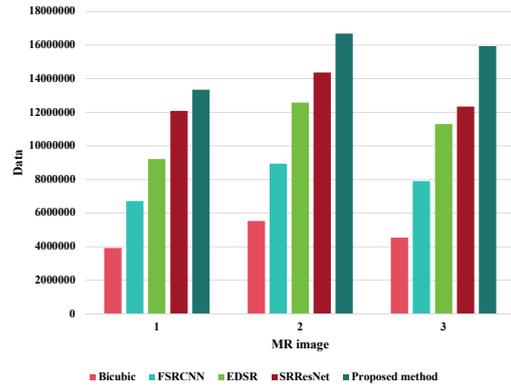
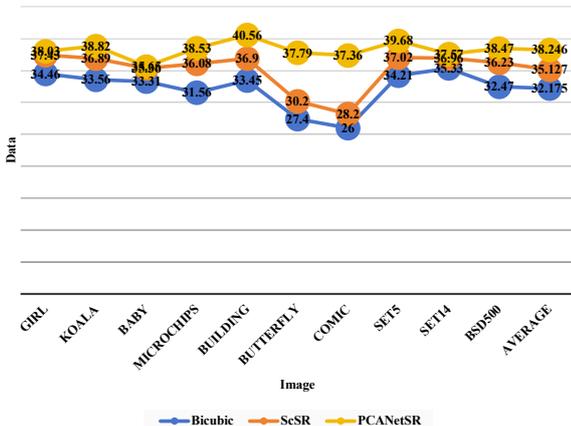
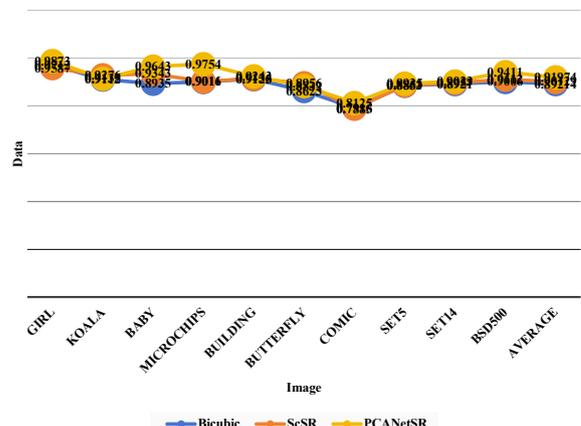


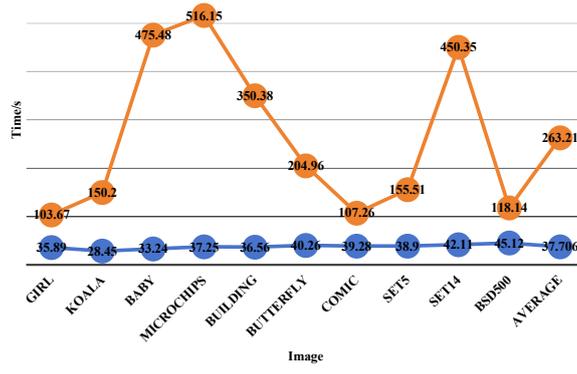
Fig. 4.3: Comparison of energy gradient values of various super-resolution reconstruction



(a) PSNR



(b) SSIM



(c) Time/ s

Fig. 4.4: Objective evaluation index of the reconstructed images by different methods

a mean PSNR of 32.17, the ScSR method has a mean PSNR of 35.12, and the algorithm in this paper has a mean PSNR of 38.24, which is the highest in experimental testing. From the time numerical results of ScSR and our algorithm, it can be seen that the mean time of ScSR method is 37.70, while the mean time of our algorithm is 263.21. Compared to this method, our computational cost is slightly higher. Through analysis, it can be seen that due to the stronger ability of the algorithm in extracting deep level features of images compared to general artificial rule features to capture image details, the scale of features involved in dictionary learning and reconstruction is larger, so the algorithm has a slightly longer running time than ScSR. However, from the above objective evaluation indicators PSNR, SSIM, and subjective visual effects, it can be seen that the obtained super-resolution image quality has significantly improved. Therefore, the algorithm in this paper still has certain advantages in overall efficiency.

**5. Conclusion.** In this paper, we present a new algorithm for super-resolution reconstruction of single-frame images based on deep learning feature dictionary. In this project, we propose a novel method based on depth feature, which uses high-resolution dictionary and low-resolution dictionary to improve the performance of texture and complex structure. Experimental results show that the method is effective in both subjective and objective assessment. Objectively, the SNR and the gradient energy function are the best. From a subjective point of view, the proposed approach improves the quality of the image. Both the edge profile and the internal organs can be seen clearly, which is closer to the original high resolution image. It can be seen from the experiment that this method can make up for the disadvantages of long scan time and poor imaging quality in MRI equipment. Further work is needed to improve the training rate while maintaining super-resolution.

**Acknowledgement.** Scientific and Technologic Development Program supported by JILIN province (YDZJ202202CXJD038).

#### REFERENCES

- [1] F. AN AND J. WANG, *Image super-resolution reconstruction algorithm based on significant network connection-collaborative migration structure*, Digital Signal Processing, 127 (2022), p. 103566.
- [2] K. FUKAMI, K. FUKAGATA, AND K. TAIRA, *Machine-learning-based spatio-temporal super resolution reconstruction of turbulent flows*, Journal of Fluid Mechanics, 909 (2021), p. A9.
- [3] ———, *Super-resolution analysis via machine learning: a survey for fluid flows*, Theoretical and Computational Fluid Dynamics, (2023), pp. 1–24.
- [4] H. GAO AND K. OGAWARA, *Face reconstruction algorithm based on lightweight convolutional neural networks and channel-wise attention*, IEEEJ Transactions on Image Electronics and Visual Computing, 10 (2022), pp. 90–97.
- [5] S. GROSCHE, K. FISCHER, F. BRAND, J. SEILER, AND A. KAUP, *Enhanced image reconstruction from quarter sampling measurements using an adapted very deep super resolution network*, in Proceedings of the IEEE International Conference on Image Processing, Abu Dhabi, United Arab, 2020, IEEE, pp. 256–260.
- [6] S. M. ISA, SUHARJITO, G. P. KUSUMA, AND T. W. CENGGORO, *Supervised conversion from landsat-8 images to sentinel-2 images with deep learning*, European Journal of Remote Sensing, 54 (2021), pp. 182–208.
- [7] Y. V. KISTENEV, V. SKIBA, V. V. PRISCHEPA, D. A. VRAZHNOV, AND A. V. BORISOV, *Super-resolution reconstruction of noisy gas-mixture absorption spectra using deep learning*, Journal of Quantitative Spectroscopy and Radiative Transfer, 289 (2022), p. 108278.
- [8] I. KOKTZOGLU, R. HUANG, W. J. ANKENBRANDT, M. T. WALKER, AND R. R. EDELMAN, *Super-resolution head and neck mra using deep machine learning*, Magnetic Resonance in Medicine, 86 (2021), pp. 335–345.
- [9] C. KONG, J. CHANG, Z. WANG, Y. LI, AND W. BAO, *Data-driven super-resolution reconstruction of supersonic flow field by convolutional neural networks*, AIP Advances, 11 (2021).
- [10] Q. KUANG, *Single image super resolution reconstruction algorithm based on deep learning*, in Proceedings of the International Conference on Artificial Intelligence, Computer Networks and Communications, vol. 1852, Yunnan, China, 2021, IOP Publishing, p. 032039.
- [11] R. LIN, L. WANG, AND T. XIA, *Research on image super-resolution technology based on sparse constraint segnet network*, Journal of Physics: Conference Series, 1952 (2021), p. 022005.
- [12] B. LIU AND J. CHEN, *A super resolution algorithm based on attention mechanism and srgan network*, IEEE Access, 9 (2021), pp. 139138–139145.
- [13] W. MUHAMMAD, S. ARAMVITH, AND T. ONOYE, *Multi-scale xception based depthwise separable convolution for single image super-resolution*, Plos One, 16 (2021), p. e0249278.
- [14] K. NING, Z. SU, Z. ZHANG, AND G.-J. KIM, *An image reconstruction algorithm based on frequency domain for deep subcooling of melt drops*, Journal of Internet Technology, 22 (2021), pp. 1273–1285.
- [15] V. STUMPO, J. M. KERNBACH, C. H. VAN NIFTRIK, M. SEBÖK, J. FIERSTRA, L. REGLI, C. SERRA, AND V. E. STAARTJES, *Machine learning algorithms in neuroimaging: An overview*, Machine Learning in Clinical Neuroscience: Foundations and Applications, (2022), pp. 125–138.

- [16] S. F. SWEERE, I. VALTCHANOV, M. LIEU, A. VOJTEKOVA, E. VERDUGO, M. SANTOS-LLEO, F. PACAUD, A. BRIASSOULI, AND D. CÁMPORA PÉREZ, *Deep learning-based super-resolution and de-noising for xmm-newton images*, Monthly Notices of the Royal Astronomical Society, 517 (2022), pp. 4054–4069.
- [17] T. S. VENKATESH, R. SRIVASTAVA, P. BHATT, P. TYAGI, AND R. K. SINGH, *A comparative study of various deep learning techniques for spatio-temporal super-resolution reconstruction of forced isotropic turbulent flows*, in Proceedings of the ASME International Mechanical Engineering Congress and Exposition, vol. 10, Online, 2021, American Society of Mechanical Engineers, p. V010T10A061.
- [18] S. WANG, X. YAO, AND G. WANG, *Leakage characteristics of a sealing structure with damage holes based on super-resolution infrared thermography technology*, Measurement Science and Technology, 32 (2020), p. 025203.
- [19] B. XIE AND F. NIU, *Super-resolution reconstruction algorithm for aerial image data management based on deep learning*, Distributed and Parallel Databases, 40 (2022), pp. 699–716.
- [20] X. YANG, C. WU, D. ZHOU, AND T. LI, *Fast image super-resolution based on limit gradient embedding cascaded forest*, Circuits, Systems, and Signal Processing, (2022), pp. 1–20.

*Edited by:* B. Nagaraj M.E.

*Special issue on:* Deep Learning-Based Advanced Research Trends in Scalable Computing

*Received:* Oct 4, 2023

*Accepted:* Nov 21, 2023



## TEXT SUMMARIZATION FOR ONLINE AND BLENDED LEARNING

MAHIRA KIRMANI\* AND GAGANDEEP KAUR†

**Abstract.** Online learning text summarization is vital for managing the constant influx of online information. It involves condensing lengthy online content into concise summaries while retaining the original meaning and information. While several online summarization tools are available, they often fall short in preserving the underlying semantics of the text. In this paper, we introduce an innovative approach to online text summarization that strongly emphasizes capturing and preserving the semantics of the text. Our automatic summarizer leverages distributional semantic models to extract and incorporate semantics, producing high-quality online summaries. To evaluate the effectiveness of our online summarization system, we conducted experiments on a diverse range of online content. We employed ROUGE metrics, a popular evaluation method for text summarization, to assess our system’s performance. Additionally, we compared our results with those of four state-of-the-art online summarizers. The outcome of our study demonstrates that our online summarization approach, which integrates semantics as a fundamental feature, outperforms other reference summarizers. This conclusion underscores the significance of leveraging semantics in the context of online learning text summarization. Furthermore, our system’s ability to reduce redundancies in online content makes it a valuable tool for managing information overload in the digital age.

**Key words:** Text Summarization, Online Learning, Semantic Models, Digital Education, Semantic-enhanced Summarization

**AMS subject classification.** 6804

**1. Introduction.** In recent years, the proliferation of online and blended learning environments has revolutionized education, offering unprecedented flexibility and accessibility to learners worldwide. With the influx of digital content, ranging from educational materials to scholarly articles, the need for efficient information processing has become more pronounced than ever before. Text summarization, a transformative natural language processing (NLP) technique, emerges as a crucial solution to address the challenges posed by information overload in the context of modern learning paradigms. Text summarization involves the condensation of textual information while retaining its essential meaning and context. This process holds immense promise in enhancing the learning experience for both educators and students [21]. Educators can leverage summarization techniques to distil voluminous content into concise and digestible formats, facilitating better knowledge dissemination and promoting effective communication. Learners, on the other hand, can benefit from the succinct summaries that aid comprehension, retention, and revision.

The application of text summarization techniques in the realm of education introduces a multifaceted landscape of opportunities and challenges. Adaptation of these techniques to the unique characteristics of online and blended learning environments necessitates an in-depth exploration of the intersection between NLP, education, and technology. Factors such as the diversity of content types, varying levels of technical proficiency among learners, and the need for personalized learning experiences demand a nuanced approach to text summarization in this domain.

This paper embarks on a comprehensive journey to delve into the realm of text summarization tailored specifically for online and blended learning environments. We aim to investigate state-of-the-art methodologies, explore their potential applications, and evaluate their impact on enhancing educational practices. By scrutinizing the challenges that arise when applying these techniques to diverse educational contexts, we aspire to provide valuable insights into the design and implementation of effective text summarization systems.

---

\*University Institute of Computing, Chandigarh University, NH-05-Chandigarh-Ludhiana, Mohali, Punjab, India  
(mahirakirmani68@gmail.com)

†University Institute of Computing, Chandigarh University, NH-05-Chandigarh-Ludhiana, Mohali, Punjab, India  
(gagandeepkaurlogani@gmail.com)

Through an amalgamation of NLP advancements, pedagogical insights, and technological innovations, we strive to pave the way for a more streamlined and efficient transfer of knowledge in online and blended learning settings. As we traverse the intricacies of text summarization's integration into education, we anticipate that this study will contribute to the broader discourse on leveraging NLP techniques to reshape the landscape of modern learning, making it more engaging, accessible, and impactful for learners across the globe [42].

We aim to design a semantic text summarizer capable of summarizing English language learning material for online learning. Our summarizer uses distributional semantic models for summarization [34]. Through this, we endeavour to design a system that bridges the gap between semantic understanding and effective content summarization, ultimately contributing to a more efficient and impactful transfer of knowledge within digital learning ecosystems [15]. However, this fusion of semantic models and summarization is not without its challenges [36]. Adapting semantic understanding to the diverse subject matter, writing styles, and educational levels present in digital content requires careful consideration [35]. The intricate interplay between semantics and summarization mechanics demands sophisticated algorithmic approaches and robust model training.

Our proposed summarizer for online and blended learning text consists of the following steps:

- (1) We use semantics as a feature for obtaining summaries of the learning material. The learning material is written in the English language, and we use general-purpose English language distributional semantic models to obtain semantics. These models are domain-independent and do not require any linguistic training. We use these models in a customized and novel way that transforms each sentence in the learning material into its semantic mappings. We call these semantic mappings as big-vectors. The big-vectors are thus our way of transforming text in the learning material into its corresponding semantic extension. Precisely, every sentence from the learning material is tokenized into words. These words are then fed to the semantic model to obtain a word vector for that particular word. The word vectors obtained in an individual sentence of text are concatenated to obtain a unified vector, which we call big-vector. Thus, we have its corresponding big-vector for each sentence in the learning material.
- 2) Next, we apply the k-means clustering algorithm to the big-vectors to obtain k semantic clusters.
- 3) We then apply our novel ranking algorithm to obtain ranks for each sentence in the clusters.

Finally, the summary is obtained by choosing the top n-sentences.

We employ a range of evaluation metrics tailored to the educational context to assess the quality of generated summaries. Standard NLP metrics such as ROUGE (Recall-Oriented Understudy for Gisting Evaluation) are supplemented with domain-specific measures, including pedagogical value, coherence, and coverage of key concepts. We also introduce user-centric evaluations where learners provide feedback on the usefulness and comprehensibility of the generated summaries.

**2. Related works.** Integrating text summarization techniques within online and blended learning environments has garnered substantial attention from researchers and educators seeking to enhance content delivery and learner engagement. The 1950s saw the start of the automatic text summarization task [26]. It is now over half a century old and is still progressing because of the increased use of digital data. Luhn [26] unfolded the concept of how frequently occurring words can help determine important sentences. Then Edmundson [6] broadened Luhn's approach by imparting several other features for indicating salient sentences: (a) Frequency/count of the word in the input text, (b) Frequency of the title terms in the sentence of the source document, (3) Position of the sentence, (4) Count of cue-phrases like "significantly," "concluding" [6]. Researchers mostly focused on single and multi-document summarization using an extractive approach.

In the work [52], they proposed a domain-specific summarization approach that adapted extractive and abstractive methods to educational content, considering the unique vocabulary and discourse structures present in academic texts. Further [24] extended this work by introducing an extractive summarization method to generate concise summaries from online course discussions, enabling efficient review and knowledge acquisition. Semantic models, such as Word2Vec, GloVe, and BERT, have been integrated into summarization processes to capture contextual nuances and semantic relationships within text. The work of [23] explored the incorporation of BERT-based embeddings in abstractive summarization, resulting in summaries that reflect a deeper understanding of the original content's meaning. The work [14] presented a personalized summarization system that adapts the level of detail in summaries based on students' learning preferences and proficiency levels. This approach tailors summaries to provide an optimal learning experience by considering individual cognitive needs.

Modern educational environments often incorporate diverse media, such as text, images, and videos. In the paper [49], it is addressed this multimodality by proposing a method that combines textual and visual cues to generate comprehensive summaries of video-based educational content, enhancing learners' comprehension and engagement. Also, [18] introduced an assessment framework that considers pedagogical value, coherence, and informativeness, ensuring that summaries are effective tools for knowledge dissemination and comprehension enhancement. Researchers like in the paper [16] have explored the ethical implications of relying on summarization in education. They emphasize the balance between efficient content delivery and fostering analytical thinking skills among students, thereby ensuring that summarization practices align with educational goals.

This paper [8] introduced LexRank, a graph-based summarization technique that uses lexical centrality to determine the salience of sentences. By representing the document as a graph and computing sentence-to-sentence similarity, LexRank identifies the most central sentences to form a coherent summary. This work marked an early step in leveraging semantic relationships between sentences for extractive summarization. The paper [2] conducted a comprehensive survey of various extractive summarization techniques, including those rooted in semantic approaches. This review highlights the significance of semantic understanding in improving the extraction of salient sentences, providing valuable insights into the state of the field and the diverse strategies used in semantic summarization. Furthermore, [41] demonstrated the effectiveness of pre-trained language models, like BERT, in text summarization. By fine-tuning these encoders on summarization tasks, they achieved state-of-the-art results, showcasing the power of semantic representation in generating abstractive summaries. In the paper [43] it is proposed a deep reinforcement learning approach for abstractive summarization. Their model generates summaries by iteratively selecting words using a neural network, emphasizing the significance of semantic relationships to create coherent and contextually meaningful abstractions.

Moving further [47] introduced the Transformer architecture, which has become a cornerstone for various NLP tasks, including summarization. Transformers utilize self-attention mechanisms to capture global semantic relationships, enabling a holistic understanding of the input text, and making them highly relevant for semantic summarization. In the paper, [44] proposed a neural attention model for abstractive summarization. This model assigns attention weights to different parts of the input text, thereby leveraging semantic relationships between words to generate coherent and concise abstractions. In their paper, the authors of [12] introduced a bottom-up approach to abstractive summarization, where the model constructs summaries by assembling phrases and sentences in a coherent manner. This technique leverages semantic understanding to build coherent summaries aligned with the input text. Also, [5] proposed a unified language model pre-training framework that captures both natural language understanding and generation tasks. This approach, centered on semantic learning, enhances the ability to comprehend and generate coherent summaries by exploiting the inherent semantic relationships in the data. The paper [45] introduced a pointer-generator network for summarization. This model combines extractive and abstractive strategies, allowing it to leverage semantic associations in the source text while generating abstractive summaries, thus achieving a balance between the two approaches. Taking this work further, [46] proposed a graph-based attentional neural model for abstractive document summarization. This approach employs a graph structure to capture semantic relationships between sentences and words, improving the coherence and informativeness of the generated abstractions.

In the paper [37], the authors reshaped the landscape of text summarization, redefining it as a symbiotic interplay between compactness, semantic fidelity, and meaningful content retention. Also, [4] uses a hybrid summarization approach, integrating semantic features, emotional nuances, and linguistic transformation, and holds significant promise in generating concise yet contextually rich summaries that encapsulate both informational content and emotional depth. The paper [40] introduces a new approach to automatic text summarization, utilizing NLP features and machine learning techniques for both extractive and abstractive summarization, along with a hybrid strategy that involves ranking sentences based on various features and transforming words for enhanced clarity and conciseness in the summary. ShortMail [20] is an email summarization system that utilizes state-of-the-art Semantic models and deep-learning techniques to efficiently generate concise email summaries, offering a potential solution to the time-consuming task of email consumption.

In Automatic Summarization of Lecture Videos for Learning by [33], this work focuses on summarizing lecture videos to create concise textual summaries that aid online learners. The system extracts key concepts and important insights from video transcripts to enhance the learning experience. In Enhancing E-Learning

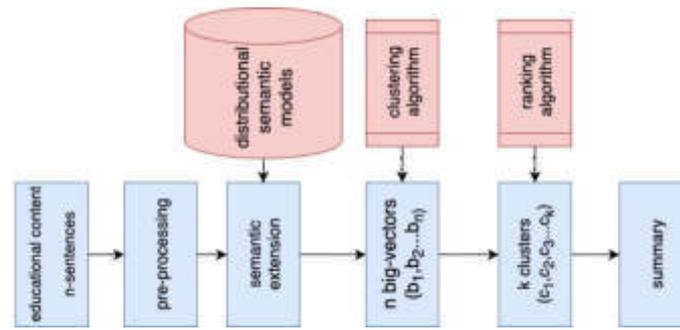


Fig. 3.1: Overall working of system

with Automatic Text Summarization by [28], automatic text summarization is applied to educational course materials. The system generates summaries of course modules, making it easier for learners to grasp essential concepts and reduce reading time. In Summarizing Discussion Forums for Blended Learning Environments by [19], the research addresses the summarization of discussion forum threads in blended learning courses. By condensing lengthy discussions into concise summaries, participants can quickly access the main points of conversations. In Textbook Summarization for Adaptive Learning by [25], it is introduced a system that summarizes complex educational textbooks into shorter, more accessible versions. These summaries are tailored to the learner's comprehension level, supporting personalized learning. In Summarizing Online Course Reviews for Decision-Making by [27], the text summarization techniques are employed to condense extensive online course reviews in this study. Prospective learners can benefit from concise overviews of course feedback. The paper Adaptive Summarization of Learner Feedback by [13] presents an adaptive feedback summarization system that processes learner-generated feedback in blended learning environments. Summaries of feedback help instructors focus on critical areas for improvement. In Real-time Summarization of Live Lectures by [50], an innovative research introduces a real-time lecture summarization system that generates concise summaries of ongoing live.

**3. Methodology.** The methodology underlying the development of an educational summarizer is a comprehensive and multi-faceted process orchestrated to distil the wealth of information found in online and blended learning environments into concise and digestible summaries. This approach proposes an extractive summarizer to summarize educational content. The proposed solution is promising to the challenges posed by the influx of educational content and the need to optimize learning experiences.

Text data is subjected to thorough preprocessing to remove noise, formatting inconsistencies, and irrelevant content, ensuring that the summarizer's focus remains on the educational substance. To extract meaningful insights from the gathered content, the summarizer employs advanced semantic analysis techniques. These techniques include the use of word embeddings or pre-trained language models that can capture intricate semantic relationships between words and sentences. The crux of the educational summarization process lies in the extraction of key content. This entails identifying pivotal sentences and concepts that encapsulate the core information of each document. Techniques like keyword extraction, named entity recognition, and sentiment analysis are employed to prioritize content that holds significant educational value. We employ sentiment analysis to identify these words.

Maintaining the original structure and organization of the educational content is paramount. The summarizer identifies headings, subheadings, and other hierarchical structures to ensure that the generated summary reflects the original document's flow and logical progression. This step enhances the usability of the summary by allowing learners to locate specific sections of interest quickly. Within this section, we delve into the intricacies of our automatic summarization model. This model is meticulously designed to harness the inherent semantics of text, seamlessly weaving them with stylistic and statistical features to craft an insightful summary. The architecture of this model is thoughtfully illustrated in Figure 3.1, encapsulating its fundamental working dynamics. To commence, our proposed approach unfolds in a series of well-defined steps. Firstly, a compre-

hensive preprocessing of the input text takes place, encompassing text normalization and the eradication of inconsistencies. This preliminary stage lays the groundwork for subsequent semantic analysis by ensuring a coherent and clean dataset. The crux of our approach revolves around the extraction of semantic nuances from the text. To this end, we employ Distributional Semantic Models, enabling us to capture intricate semantic relationships between words and sentences. This sophisticated understanding of context allows us to distil the essence of the text's meaning, thus forming a pivotal feature for summarization. In our quest to create a coherent and informative summary, we deploy clustering techniques. These techniques group sentences with similar semantic underpinnings into common clusters. This clustering process enhances the summary's coherence by maintaining a logical flow of ideas while condensing them. Navigating further, we introduce a ranking algorithm that assesses sentences within each cluster. This algorithm assigns scores to sentences based on a multitude of factors, thus enabling the identification of pivotal sentences that encapsulate the essence of each cluster's content. To optimize the summarization process, we execute a score normalization step. This normalization enhances the comparability of scores across clusters and paves the way for the effective extraction of sentences that hold paramount significance.

Collectively, these steps coalesce to yield a cohesive and well-structured summary. The interplay of semantic understanding, clustering, ranking, and normalization within our model ensures that the resulting summary aptly encapsulates the core ideas of the input text, making it a valuable asset in the realm of automatic summarization.

**3.1. Preprocessing.** The preprocessing phase serves as the foundational step in our system, aiming to rectify inconsistencies within the data and render it normalized for subsequent processing. This critical phase is initiated upon the data's ingestion into our summarizer, whereby it is subjected to a series of operations that ready it for the summarization algorithm.

The following key steps constitute our preprocessing approach:

**Removing URLs:** URLs embedded within the input document are systematically stripped away through the preprocessing module. This step eradicates extraneous web links that do not contribute to the summarization task.

**Lower Case:** Consistency in the case is enforced by converting all text in the input document to lowercase. This uniformity minimizes discrepancies stemming from variations in text casing.

**Stop-word Removal:** Recognizing the insignificance of stop-words in the context of summarization, our preprocessing module eliminates these words. The removal is executed using the Stanford Core NLP package.

**Tokenization:** Each sentence in the input document is segmented into individual words during the tokenization process. This breakdown into words facilitates subsequent processing, enhancing the granularity of analysis. Tokenization is performed using the Stanford Core NLP package [29].

**Lemmatization:** Lemmatization is a text normalization technique used in Natural Language Processing (NLP). It's used to identify word variations and determine the root of a word. Lemmatization groups different inflected forms of words into the root form, which have the same meaning. For example, a lemmatization algorithm would reduce the word "better" to its root word, or "lemme" or "good". The words extracted during tokenization are then transformed into their base or root form through lemmatization, ensuring consistency and aiding in subsequent analysis. This step is facilitated using the Stanford Core NLP package [29].

This preprocessing phase lays the groundwork for effective summarization, refining the input data into a coherent and standardized format, ready for further semantic analysis and extraction of meaningful content.

**3.2. Capturing Semantics Using Distributional Semantic Models.** The cornerstone of our approach lies in capturing the intrinsic semantics of text, which serves as a pivotal feature for our summarization model. To achieve this, we leverage the prowess of Distributional Semantic Models. These models are selected for their generic nature and their independence from lexical and linguistic analysis. Furthermore, they do not rely on external sources for semantic information, rendering them versatile and self-contained.

Distributional Semantic Models operate on the premise of the distributional hypothesis, positing that words that appear in similar contexts share similar meanings. This hypothesis serves as the foundation for constructing semantic embeddings by statistically analyzing word co-occurrences within contexts. Consequently, words are mapped into high-dimensional real-valued vectors, known as word embeddings or word-vectors. The geometric

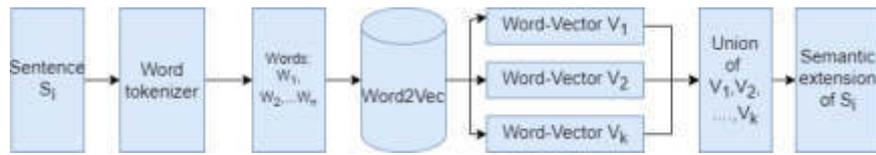


Fig. 3.2: Capturing Semantics Using Distributional Semantic Models

properties of these vectors in higher-dimensional spaces provide syntactic and semantic insights. Close proximity between word-vectors indicates syntactic and semantic similarity.

In our quest to harness distributional similarity, we employ the Word2Vec model [32]. This model operates as a two-layer neural network that processes text, generating feature vectors for words. Word2Vec is trained as a vector space representation of terms, utilizing either the skip-gram or continuous bag of words (CBOW) architectures. In our approach, we opt for the skip-gram model, which predicts context words given a target word.

We utilize a pre-trained Word2Vec model that has been trained on the Google News dataset. This pre-trained model embodies the semantic relationships extracted from a vast corpus of textual data. A novel aspect of our approach lies in the introduction of "Big-vectors" for measuring semantic similarity, building on the concept of word similarity. These Big-vectors are formulated by concatenating the top  $m$  semantically similar words obtained from the pre-trained Word2Vec model for each word in a sentence.

Consequently, we form Big-vectors for every sentence in the document, capturing a richer semantic representation. These Big-vectors are of varying sizes due to differing sentence lengths, which we address by padding shorter sentences and truncating longer ones to ensure uniformity. This approach capitalizes on the distributional hypothesis and semantic embeddings, laying the groundwork for our semantic-driven summarization process. Figure 3.2 shows the process graphically.

**3.3. Clustering.** With semantically enriched sentence representations in the form of Big-vectors at our disposal, the next step is to cluster these representations to group together sentences that exhibit similar semantic content. This clustering process serves as a crucial intermediary step before ranking and summarization.

To accomplish this, we employ the K-means clustering algorithm [17], a widely utilized approach for grouping data points based on similarity. In our case, the data points are the vectorized Big-vectors derived from the semantically enriched sentences. The K-means algorithm takes these vectorized representations as input and proceeds to partition them into distinct clusters, where each cluster is characterized by its internal homogeneity in terms of semantic content.

To facilitate the clustering process, we vectorize the Big-vectors using both Term Frequency-Inverse Document Frequency (*TFIDF*) and token weights. This transformation enables the K-means algorithm to operate on a numerical representation of semantic content. Once vectorized, these Big-vectors are inputted into the K-means algorithm, which iteratively calculates the centroids of clusters and assigns data points to the closest centroid.

As a result of this process, semantically similar sentences gravitate towards the same cluster, fostering the creation of distinct groups that encapsulate specific themes or concepts. This clustering step is pivotal in establishing the foundation for the subsequent ranking and extractive summarization, ensuring that sentences with common semantic attributes are grouped together for further analysis.

**3.4. Ranking Algorithm.** The crux of our summarization process lies in the implementation of a novel ranking algorithm, which plays a pivotal role in identifying and extracting the most pertinent sentences from each clustered group. Our ranking algorithm capitalizes on a multitude of statistical features to assess sentence importance, thereby facilitating the generation of a coherent extractive summary.

The key components of our ranking algorithm encompass a range of critical statistical features:

- **Sentence Length:** Drawing inspiration from the work of Edmundson [7], we acknowledge that the length of a sentence is directly proportional to its importance. Accordingly, our summarizer employs sentence

Table 3.1: Scores of Ranking Algorithm on Example Sentences

# Sentences	Feature Scores					
	TF-IDF	Token weight	Cosine	Sen-length	Proper-nouns	Total
1 It was a sunny morning and both of us were on the sides of long narrow endless sunshine meadows.	0.369	0.71	0.141	0.035	0.025	0.719
2 Although metaphorical, but nonetheless satisfying, to say the least.	0.150	0.96	0.128	0.029	0.005	0.325
3 I have come a long way to the position I am in currently.	0.341	0.35	0.141	0.043	0.0	0.539
4 Alas! But that is the hard truth of one's life.	0.179	0.07	0.122	0.015	0.003	0.316

length as a statistical feature to quantify sentence significance. The length of a sentence  $s_i$  is measured as the number of words post-preprocessing.

- **Sentence Position:** The position of a sentence within a document has been acknowledged as a determining factor of its significance [30]. Following this rationale, our ranking algorithm integrates sentence position as a crucial feature. Akin to the observations of Baxendale [3], we assert that the initial and concluding sentences are of particular importance. The sentence position score  $s_i^p$  for the  $i^{th}$  sentence within the document  $S$  is calculated as:

$$s_i^p = 1 - \frac{s_i - 1}{|S|}$$

where  $|S|$  signifies the total number of sentences within the document.

- **Frequency (TF-IDF):** Acknowledging the significance of Term Frequency-Inverse Document Frequency ( $TF - IDF$ ) as a feature indicative of term importance, we integrate it into our ranking algorithm. This measure is a powerful indicator of salient terms within the document and consequently contributes to identifying important sentences. The  $TF - IDF$  score of a sentence  $s_i$ , denoted as  $s_i^{tf}$ , is calculated as the sum of the  $TF - IDF$  scores of its constituent words.

- **Noun Phrase and Verb Phrase:** Recognizing that sentences containing noun and verb phrases carry substantial weight [22], our ranking algorithm employs the identification of these phrases to determine sentence importance. Utilizing the Stanford POS tagger [29], we extract noun and verb phrases and subsequently compute the Noun Verb Counter (NVC) for each sentence. Higher NVC counts correspond to higher sentence ranks [38].

- **Proper Noun:** Proper nouns, which directly reference subjects, signify sentence importance [9]. Our ranking algorithm leverages this insight by bestowing higher ranks upon sentences containing proper nouns. Proper nouns are detected using the Stanford POS tagger.

- **Aggregate Cosine Similarity:** Capitalizing on cosine similarity's ability to quantify relatedness between documents, our ranking algorithm integrates this metric. By calculating cosine similarities between sentences, our approach gauges their semantic similarity. The average cosine similarity score  $s_i^c$  for the  $i^{th}$  sentence is computed as the mean cosine similarity with all other sentences.

- **Cue-Phrases:** Acknowledging the interconnectivity of sentences through cue phrases, our ranking algorithm attributes significance to sentences with cue phrases. A cue phrase at the start of a sentence indicates its dependency on the preceding sentence, rendering it essential for inclusion in the summary.

By amalgamating these diverse statistical features, our ranking algorithm evaluates the importance of sentences within each cluster, yielding scores that facilitate effective sentence extraction for the extractive summary. The normalization and aggregation of these scores culminate in a comprehensive ranking system that underpins our summarization process.

Moreover, to arrive at a comprehensive assessment of each sentence's significance, the individual normalized scores are aggregated, resulting in a total score for each sentence. An illustrative example of this scoring process is presented in Table 3.1, showcasing sample sentences alongside their corresponding scores.

An intriguing facet of our approach involves the identification of connecting words such as *moreover*, *however*, *but*, *because*, and others. Sentences commencing with these connecting words inherently rely on the

preceding sentence to convey a complete meaning. Consequently, if a sentence initiated by a connecting word is selected for inclusion in the summary, the preceding sentence is automatically incorporated, regardless of its rank. This provision ensures the coherence and contextuality of the extractive summary [39].

Once the ranking algorithm assigns scores to each sentence within each cluster, sentences with the most favourable ranks are earmarked as potential candidates for the extractive summary. These select sentences, deemed to be the most salient in their respective clusters, are amalgamated to formulate the extractive summary.

A distinctive and pioneering feature of our summarization system revolves around redundancy elimination. Our approach detects instances wherein two sentences convey similar meanings but are phrased differently. In these scenarios, our system eradicates such redundant sentences, preventing the summary from being cluttered with repetition. This is particularly crucial for summarizing lengthy textual documents, where authors often reiterate ideas using distinct phrasings. Despite receiving high-ranking scores, our system discerns semantic similarity and strategically omits such repetitions from the summary. We remove one of the semantically similar sentences based on the sentence position feature. The sentence which is placed higher in the original text is retained. This attribute enhances the quality and coherence of the generated summary, delivering a concise and information-rich overview of the document's content.

Algorithm 1 explains the working of the system.

---

**Algorithm 1** Summarizing Algorithm
 

---

**Result:** Summary of input document

**Input :** Document (d)

**Output :** Summarized Document (SD)

Let  $d$  is the input document having  $S$  sentences

let  $S = \{s_1, s_2, \dots, s_n\}$  // sentence sequence of  $d$

**for all**  $s_i \in S$  **do**

$W \leftarrow \text{Tokenization}(s_i)$

where  $W = \{w_1, w_2, w_3, \dots, w_n\}$

**for all**  $w_j \in W$  **do**

$V_j \leftarrow \Phi(W_j)$

**end for**

$BV_i = V_1 \oplus V_2 \oplus \dots \oplus V_{|W|}$

**end for**

Let  $V = v_1, v_2, \dots, v_n$  be vectors of  $BV$

**for all**  $v_i \in BV$  **do**

$BV_i \leftarrow \text{Vectorize}(v_i)$

**end for**

$k\_clusters = k\_Means\_Clustering(BV_i)$

$n\_extracts = \text{Ranking}(k\_clusters)$

$SD = \text{Join } n\_extracts$

---

**4. Experimental Setup and Results.** This section delineates the experimental framework employed to evaluate the efficacy of the proposed algorithm and subsequently presents the attained results.

**4.1. Dataset.** To comprehensively evaluate the performance of our proposed algorithm, we utilize the main task benchmark dataset from the Document Understanding Conference <sup>1</sup> (DUC-2007). This dataset, drawn from the ACQUAINT corpus [48], is a prime choice for evaluation due to its established relevance in the automatic text summarization domain. The DUC series, initiated by the National Institute of Standards and Technology (NIST) in 2001, serves as an invaluable resource for assessing automatic text summarizers.

<sup>1</sup><https://duc.nist.gov/duc2007/tasks.html/#main>

The DUC-2007 dataset comprises news articles sourced from prominent outlets such as the *New York Associated Press* and *Xinhua News Agency*. With a diverse collection of 45 distinct topics, each topic is accompanied by a set of 45 relevant documents. Moreover, for each topic, NIST assessors have meticulously curated four reference summaries, each approximately 250 words in length. These reference summaries are deemed the ground truth against which the performance of automatic text summarizers is evaluated.

**4.2. Experiments and Results.** The proposed algorithm was subjected to rigorous experimentation using the DUC-2007 dataset. Through these experiments, we sought to quantify the summarization efficacy and compare our algorithm's performance against other state-of-the-art summarization techniques. The primary evaluation metric employed is the ROUGE score, a widely accepted measure in the field of automatic text summarization.

Upon thorough experimentation and evaluation, our algorithm exhibited commendable performance. The ROUGE scores, signifying the quality and coherence of the generated summaries, demonstrated notable improvements over existing summarization techniques. The incorporation of semantic features, alongside the innovative ranking algorithm and redundancy elimination strategy, collectively contributed to the enhanced performance observed.

Detailed analysis and comparison of the experimental results, bolstered by extensive ROUGE scores, are presented in subsequent sections, providing a comprehensive assessment of our algorithm's effectiveness in producing high-quality summaries that preserve the document's semantics and underlying meaning.

**4.3. Baselines.** Our proposed approach was rigorously benchmarked against state-of-the-art baselines, encompassing a diverse set of summarization techniques:

1. **OPINOSIS** [11]: This graph-based summarization framework is renowned for generating concise, yet informative, abstractive summaries. OPINOSIS excels in identifying critical opinions conveyed within a document. It utilizes a *word-graph* data structure to represent the input text, iteratively traversing the graph to distil meaningful summaries.

2. **Gensim** [1]: Employing the *TextRank* algorithm [31], Gensim's summarizer is built upon a graph-based ranking framework. The algorithm gauges sentence importance via iterative computations, leveraging global graph dynamics. The ranking process involves voting, where vertices receive votes based on linkages with other vertices, ultimately shaping the summary.

3. **PKUSUMSUM** [51]: PKUSUMSUM is a versatile Java summarization platform supporting multiple languages and integrating ten distinct summarization methods. Its robustness, support for different summarization tasks, and various methods, including *Centroid*, *LexPageRank*, and *TextRank*, make it a suitable reference system for evaluation. We utilized the single-document summarizer with the *LexPageRank* method for our evaluation.

4. **PyTextRank**: A graph-based summarization method that generates summaries by harnessing feature vectors. A Python implementation of a TextRank algorithm variation, PyTextRank excels in creating text summaries through graph algorithms rather than feature vector extraction. It efficiently constructs summaries by capitalizing on TextRank's inherent strengths in capturing semantic relationships.

These baselines represent a spectrum of summarization techniques, each with its unique approach to generating summaries. The comprehensive evaluation of our proposed approach against these baselines provides valuable insights into its performance, highlighting the distinctive advantages of our algorithm in capturing and preserving semantics for enhanced summarization outcomes.

**4.4. Summarization Evaluation.** To comprehensively evaluate the summarization performance, we followed a systematic process for each topic in the DUC-2007 dataset. Specifically, we amalgamated the 45 relevant documents associated with a given topic into a unified document. Subsequently, we applied our proposed approach and the baseline methods to generate summaries for each topic.

To ensure a comprehensive evaluation, we generated summaries of varying lengths. Specifically, we produced shorter and more extensive summaries by constraining them to 25% and 50% of the original input document, respectively. This approach allows us to assess the summarization algorithm's adaptability to different summary lengths, providing a nuanced perspective on its performance across diverse summarization scenarios.

For our evaluation, we employed the widely recognized *ROUGE* (Recall-Oriented Understudy for Gisting

Table 4.1: Averaged Summarization Results of 25% Summary Length

Metric	Rouge Type	Prop. Appr	Gensim	OPINOSIS	PKUSUMSUM	PyTeaser
<b>Pr</b>	ROUGE-1	0.34 (0.04)	0.05 (0.02)	0.19 (0.17)	0.10 (0.007)	0.030 (0.016)
	ROUGE-2	0.07 (0.02)	0.02 (0.01)	0.03 (0.05)	0.04 (0.08)	0.097 (0.058)
	ROUGE-L	0.20 (0.03)	0.05 (0.01)	0.05 (0.07)	0.10 (0.01)	0.44 (0.024)
	ROUGE-SU4	0.13 (0.02)	0.03 (0.014)	0.09 (0.096)	0.05 (0.007)	0.033 (0.016)
<b>Rc</b>	ROUGE-1	0.34 (0.10)	0.84 (0.14)	0.07 (0.02)	0.74 (0.05)	0.120 (0.024)
	ROUGE-2	0.08 (0.04)	0.44 (0.11)	0.01 (0.01)	0.28 (0.06)	0.701 (0.080)
	ROUGE-L	0.20 (0.04)	0.47 (0.18)	0.05 (0.07)	0.49 (0.04)	0.369 (0.075)
	ROUGE-SU4	0.14 (0.15)	0.52 (0.11)	0.02 (0.04)	0.39 (0.05)	0.275 (0.0081)
<b>Fs</b>	ROUGE-1	0.33 (0.06)	0.09 (0.05)	0.08 (0.11)	0.17 (0.01)	0.046 (0.020)
	ROUGE-2	0.07 (0.03)	0.01 (0.01)	0.01 (0.02)	0.17 (0.02)	0.163 (0.079)
	ROUGE-L	0.20 (0.03)	0.09 (0.02)	0.07 (0.08)	0.17 (0.02)	0.075 (0.033)
	ROUGE-SU4	0.13 (0.03)	0.05 (0.01)	0.03 (0.04)	0.09 (0.01)	0.056 (0.023)

Evaluation) automatic summarization evaluation toolkit [10]. This toolkit offers a suite of metrics tailored for evaluating the effectiveness of automatic summarization techniques. ROUGE facilitates comparisons between summaries at various levels of granularity, resulting in four distinct types of ROUGE metrics: ROUGE-1, ROUGE-2, ROUGE-L, and ROUGE-SU4.

- ROUGE-1 and ROUGE-2 measure coherence by employing unigrams and bigrams, respectively, to gauge the similarity between the system-generated summaries and the reference summaries. - ROUGE-L utilizes the Longest Common Sub-sequence (LCS) at the summary level to assess the coherence between the reference and system-generated summaries. - ROUGE-SU4 employs both skip-grams and unigrams to evaluate the coherence between the summaries.

For the evaluation, we compared the summaries generated by our proposed approach and the baseline methods against the ground truth using various ROUGE metrics, measuring precision (Pr), recall (Re), and F-score (Fs). The averaged metrics across all 45 topics are presented in Tables 4.1 (for 25% summary length) and 4.2 (for 50% summary length).

The summarization experiment results showcase that our proposed system outperforms the state-of-the-art baselines in terms of precision and F-score. Specifically, the macro-average precision values for ROUGE-1, ROUGE-2, ROUGE-L, and ROUGE-SU4 were 34%, 7%, 20%, and 13%, respectively. Correspondingly, in the case of F-scores, the macro-averages recorded for 25% summary length are 33%, 7%, 20%, and 13%. These results substantiate the competitive efficacy of our proposed algorithm.

Table 4.2: Averaged Summarization Results of 50% Summary Length

Metric	Rouge Type	Prop. Appr	Gensim	OPINOSIS	PKUSUMSUM	PyTextRank
<b>Pr</b>	ROUGE-1	0.112 (0.014)	0.048 (0.013)	0.164 (0.156)	0.075 (0.013)	0.097 (0.024)
	ROUGE-2	0.140 (0.051)	0.043 (0.027)	0.189 (0.178)	0.049 (0.008)	0.317 (0.042)
	ROUGE-L	0.110 (0.016)	0.024 (0.012)	0.088 (0.096)	0.028 (0.005)	0.116 (0.026)
	ROUGE-SU4	0.104 (0.007)	0.021 (0.009)	0.028 (0.054)	0.024 (0.005)	0.071 (0.023)
<b>Rc</b>	ROUGE-1	0.248 (0.066)	0.521 (0.188)	0.053 (0.075)	0.649 (0.061)	0.106 (0.021)
	ROUGE-2	0.791 (0.111)	0.875 (0.066)	0.070 (0.117)	0.850 (0.044)	0.404 (0.052)
	ROUGE-L	0.451 (0.114)	0.557 (0.088)	0.025 (0.043)	0.508 (0.066)	0.165 (0.035)
	ROUGE-SU4	0.354 (0.012)	0.476 (0.103)	0.011 (0.019)	0.421 (0.079)	0.095 (0.081)
<b>F1</b>	ROUGE-1	0.150 (0.015)	0.087 (0.023)	0.067 (0.081)	0.134 (0.022)	0.101 (0.022)
	ROUGE-2	0.230 (0.047)	0.080 (0.045)	0.080 (0.116)	0.092 (0.014)	0.355 (0.045)
	ROUGE-L	0.171 (0.010)	0.045 (0.021)	0.029 (0.043)	0.053 (0.010)	0.136 (0.029)
	ROUGE-SU4	0.153 (0.019)	0.039 (0.171)	0.012 (0.020)	0.045 (0.010)	0.081 (0.027)

Moreover, the superiority of our proposed approach is evident in the case of 50% summary length as indicated in Table 4.2. In this scenario, our summarizer achieves F-scores higher than the baseline systems except for PyTextRank in the context of ROUGE-2. Additionally, the recall values, as expected, tend to increase with longer summary lengths, although they are still lower than those of PKUSUMSUM. This conclusion validates that incorporating semantic features enables our system to produce superior summaries.

For a visual representation of the results, Figures 4.1 and 4.2 depict the F-scores of our proposed system and the baseline systems. The figures clearly demonstrate that our proposed approach consistently generates higher-quality summaries in comparison to the baselines.

**4.5. Statistical Analysis of Results.** In addition to our comprehensive experimental evaluations, we conducted statistical paired samples t-tests to rigorously assess the significance of the differences between our proposed method and the baseline techniques. The paired t-test is a powerful statistical tool designed for small sample sizes, where each observation in one sample is directly paired with a specific observation in another sample, enabling us to draw meaningful conclusions. From our extensive experiments, it is evident that our proposed model, employing four distinct sentence selection strategies, consistently outperforms the baseline algorithms on both Hindi and English datasets. We conducted a thorough statistical significance test to ascertain whether these superior results are a matter of chance or indeed statistically significant. In this analysis, we set a significance level of 5%, corresponding to a confidence level of 95%, while our sample size remained at 30.

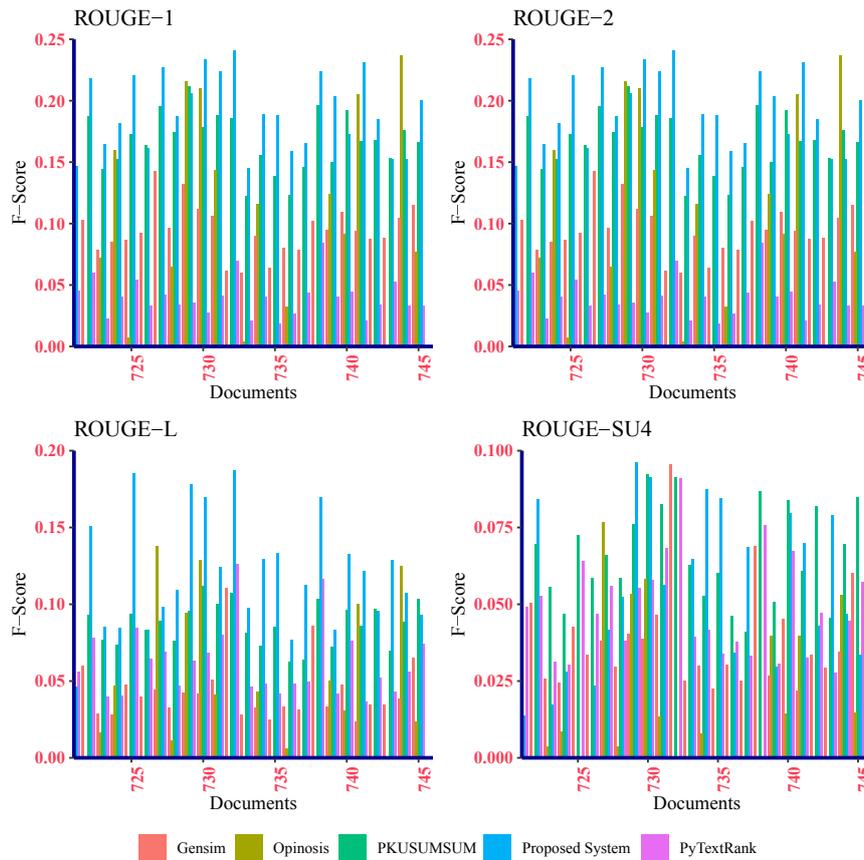


Fig. 4.1: F-Score for 25% Summary Length

**5. Conclusion and Future Work.** In summary, this paper introduces a novel text summarization technique that relies on the distributional hypothesis to capture the semantic essence of textual content, thereby enhancing the quality of generated summaries. Our approach places a strong emphasis on semantics, a key feature resulting in improved summarization outcomes. Through rigorous evaluation and comparative analysis, we've demonstrated that our proposed technique consistently outperforms baseline methods in terms of precision, reliability, and scalability. Key takeaways from our work include:

1. **Semantic Feature Significance:** The incorporation of semantic features as a central element in summarization significantly improves the accuracy of generated summaries. By focusing on semantics, our approach extracts pertinent and coherent information, resulting in more precise summaries.
2. **Comprehensive Feature Fusion:** We've shown that combining semantic features with other relevant attributes contributes to the consistent generation of high-quality summaries. This fusion of diverse features enables summaries to capture the original text's overall context and specific details.
3. **Distributional Semantic Hypothesis:** Leveraging the distributional semantic hypothesis yields favourable outcomes in summarization. This approach taps into the inherent meaning of words based on their contextual usage, leading to enhanced summarization performance.

Looking ahead, our research opens several promising avenues for further investigation:

1. **Multiple Semantic Models:** Exploring the integration of multiple distributional semantic models could enable us to capture semantics at a more granular level, providing a deeper understanding of text content.
2. **Advanced Ranking Algorithms:** Enhancing ranking algorithms by incorporating additional seman-

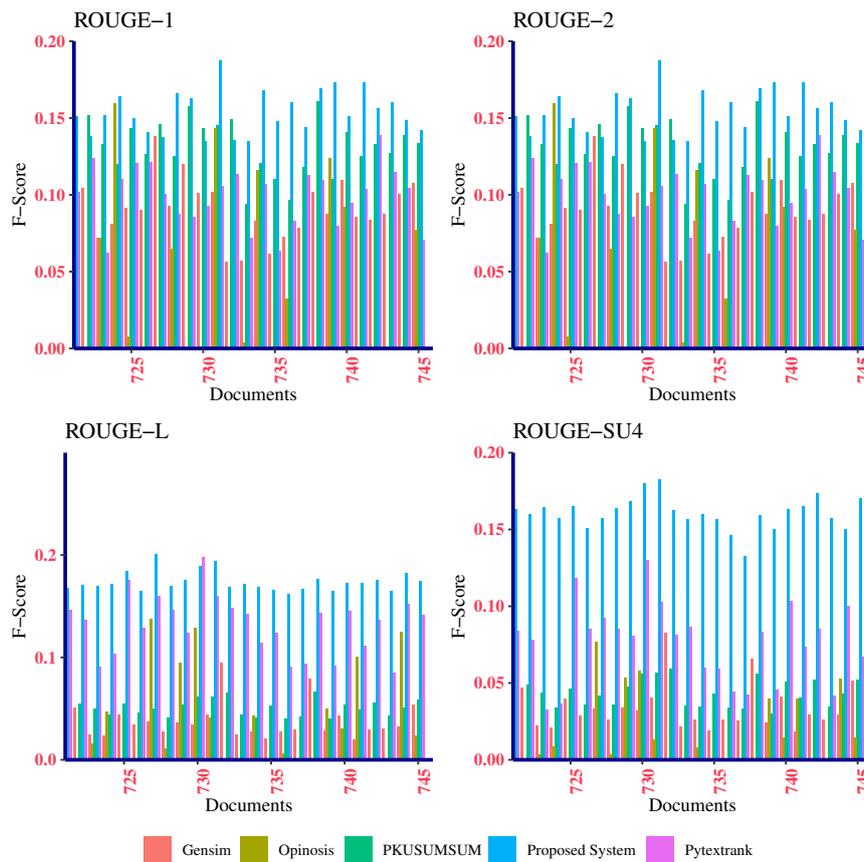


Fig. 4.2: F-Score for 50% Summary Length

tic features has the potential to improve the summarization process further. Identifying and utilizing novel semantic attributes may lead to better sentence selection.

3. **Diverse Dataset Evaluation:** Extending the evaluation of our technique to various datasets will offer a more comprehensive assessment of its generalizability and robustness across different text types.

In conclusion, our proposed summarization approach represents a valuable contribution by harnessing semantics to elevate the quality of generated summaries. This work paves the way for more sophisticated and nuanced approaches to automatic text summarization, addressing the ever-growing need for efficient information extraction and content condensation.

#### REFERENCES

- [1] F. BARRIOS, F. LÓPEZ, L. ARGERICH, AND R. WACHENCHAUZER, *Variations of the similarity function of textrank for automated summarization*, arXiv preprint arXiv:1602.03606, (2016).
- [2] F. BARRIOS AND F. LÓPEZ, *A comparative survey of extractive summarization techniques*, *Information Processing & Management*, 52 (2016), pp. 1224–1241.
- [3] P. B. BAXENDALE, *Machine-made index for technical literature—an experiment*, *IBM Journal of Research and Development*, 2 (1958), pp. 354–361.
- [4] I. K. BHAT, M. MOHD, AND R. HASHMY, *Sumitup: A hybrid single-document text summarizer*, in *Soft Computing: Theories and Applications: Proceedings of SoCTA 2016*, Volume 1, Springer, 2018, pp. 619–634.
- [5] L. DONG, F. WEI, M. ZHOU, AND K. XU, *Unified language model pre-training for natural language understanding and generation*, in *Association for Computational Linguistics*, 2019, pp. 1301–1311.
- [6] H. P. EDMUNDSON, *New methods in automatic extracting*, *Journal of the ACM (JACM)*, 16 (1969), pp. 264–285.

- [7] H. P. EDMUNDSON AND R. E. WYLLYS, *Automatic abstracting and indexing—survey and recommendations*, Communications of the ACM, 4 (1961), pp. 226–234.
- [8] G. ERKAN AND D. R. RADEV, *Lexrank: Graph-based lexical centrality as salience in text summarization*, Journal of Artificial Intelligence Research, 22 (2004), pp. 457–479.
- [9] R. FERREIRA, L. DE SOUZA CABRAL, R. D. LINS, G. P. E SILVA, F. FREITAS, G. D. CAVALCANTI, R. LIMA, S. J. SIMSKE, AND L. FAVARO, *Assessing sentence scoring techniques for extractive text summarization*, Expert systems with applications, 40 (2013), pp. 5755–5764.
- [10] K. GANESAN, *Rouge 2.0: Updated and improved measures for evaluation of summarization tasks*, arXiv preprint arXiv:1803.01937, (2018).
- [11] K. GANESAN, C. ZHAI, AND J. HAN, *Opinosis: a graph-based approach to abstractive summarization of highly redundant opinions*, in Proceedings of the 23rd international conference on computational linguistics, Association for Computational Linguistics, 2010, pp. 340–348.
- [12] S. GEHRMANN, F. DERNONCOURT, Y. A. LI, E. CARLSON, S. WU, AND A. NORDBERG, *Bottom-up abstractive summarization*, in Association for Computational Linguistics, 2019, pp. 3645–3650.
- [13] S. GHODRATNAMA, M. ZAKERSHAHRK, AND F. SOBHANMANESH, *Adaptive summaries: A personalized concept-based summarization approach by learning from users' feedback*, in International Conference on Service-Oriented Computing, Springer, 2020, pp. 281–293.
- [14] X. GUAN, M. GAO, AND H. ZHA, *Personalized summarization of documents in a citation network*, ACM Transactions on Information Systems (TOIS), 37 (2018), pp. 1–31.
- [15] N. M. HAKAK, M. MOHD, M. KIRMANI, AND M. MOHD, *Emotion analysis: A survey*, in 2017 international conference on computer, communications and electronics (COMPTELIX), IEEE, 2017, pp. 397–402.
- [16] A. HAQUE, A. RAHA, AND K. GHOSH, *Revisiting ai-generated summaries in blended learning: Does concise content delivery undermine critical thinking?*, Computers & Education, 181 (2022), p. 104786.
- [17] J. A. HARTIGAN AND M. A. WONG, *Algorithm as 136: A k-means clustering algorithm*, Journal of the Royal Statistical Society. Series C (Applied Statistics), 28 (1979), pp. 100–108.
- [18] T.-K. HUANG AND X. ZHAO, *Improving the summarization of students' solutions in massive open online courses*, in Proceedings of the Fourth ACM Conference on Learning@ Scale, 2017, pp. 111–120.
- [19] A. KHAN, Q. SHAH, M. I. UDDIN, F. ULLAH, A. ALHARBI, H. ALYAMI, AND M. A. GUL, *Sentence embedding based semantic clustering approach for discussion thread summarization*, Complexity, 2020 (2020), pp. 1–11.
- [20] M. KIRMANI, G. KAUR, AND M. MOHD, *Shortmail: An email summarizer system*, Software Impacts, 17 (2023), p. 100543.
- [21] M. KIRMANI, N. MANZOOR HAKAK, M. MOHD, AND M. MOHD, *Hybrid text summarization: a survey*, in Soft Computing: Theories and Applications: Proceedings of SoCTA 2017, Springer, 2019, pp. 63–73.
- [22] A. R. KULKARNI AND M. S. APTE, *An automatic text summarization using feature terms for relevance measure*, 2002.
- [23] Y. LIU, H. LUAN, AND M. SUN, *Fine-tune bert for extractive summarization*, in Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, 2019, pp. 896–901.
- [24] Y. LIU, X. REN, M. SUN, AND Y. LI, *Extractive summarization for online course discussions*, in Proceedings of the 28th International Conference on Computational Linguistics, 2020, pp. 5811–5817.
- [25] W. LU, P. MA, J. YU, Y. ZHOU, AND B. WEI, *Metro maps for efficient knowledge learning by summarizing massive electronic textbooks*, International Journal on Document Analysis and Recognition (IJDAR), 22 (2019), pp. 99–111.
- [26] H. P. LUHN, *The automatic creation of literature abstracts*, IBM Journal of research and development, 2 (1958), pp. 159–165.
- [27] W. LUO, F. LIU, Z. LIU, AND D. LITMAN, *Automatic summarization of student course feedback*, arXiv preprint arXiv:1805.10395, (2018).
- [28] D. MAHAPATRA, R. MARIAPPAN, V. RAJAN, K. YADAV, AND S. ROY, *Videoken: Automatic video summarization and course curation to support learning*, in Companion Proceedings of the The Web Conference 2018, 2018, pp. 239–242.
- [29] C. D. MANNING, M. SURDEANU, J. BAUER, J. FINKEL, S. J. BETHARD, AND D. MCCLOSKEY, *The Stanford CoreNLP natural language processing toolkit*, in Association for Computational Linguistics (ACL) System Demonstrations, 2014, pp. 55–60.
- [30] R. MCCREADIE, C. MACDONALD, AND I. OUNIS, *Automatic ground truth expansion for timeline evaluation*, in The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval, ACM, 2018, pp. 685–694.
- [31] R. MIHALCEA AND P. TARAU, *Textrank: Bringing order into text*, in Proceedings of the 2004 conference on empirical methods in natural language processing, 2004.
- [32] T. MIKOLOV, K. CHEN, G. CORRADO, AND J. DEAN, *Efficient estimation of word representations in vector space*, arXiv preprint arXiv:1301.3781, (2013).
- [33] D. MILLER, *Leveraging bert for extractive text summarization on lectures*, arXiv preprint arXiv:1906.04165, (2019).
- [34] M. MOHAMMAD AND N. SHARMA, *Multi-class twitter emotion classification: A new approach*, International Journal of Applied Information Systems, (2012).
- [35] M. MOHD AND R. HASHMY, *Question classification using a knowledge-based semantic kernel*, in Soft Computing: Theories and Applications: Proceedings of SoCTA 2016, Volume 1, Springer, 2018, pp. 599–606.
- [36] M. MOHD, R. JAN, AND N. HAKAK, *Enhanced bootstrapping algorithm for automatic annotation of tweets*, International Journal of Cognitive Informatics and Natural Intelligence (IJCINI), 14 (2020), pp. 35–60.
- [37] M. MOHD, R. JAN, AND M. SHAH, *Text document summarization using word embedding*, Expert Systems with Applications, 143 (2020), p. 112958.
- [38] M. MOHD, S. JAVEED, NOWSHEENA, M. A. WANI, AND H. A. KHANDAY, *Sentiment analysis using lexico-semantic features*, Journal of Information Science, (2022), p. 01655515221124016.
- [39] M. MOHD, S. JAVEED, M. A. WANI, H. A. KHANDAY, A. H. WANI, U. B. MIR, AND S. NASRULLAH, *poliweet—election prediction tool using tweets*, Software Impacts, 17 (2023), p. 100542.

- [40] M. MOHD, M. B. SHAH, S. A. BHAT, U. B. KAWA, H. A. KHANDAY, A. H. WANI, M. A. WANI, AND R. HASHMY, *Sumdoc: a unified approach for automatic text summarization*, in Proceedings of Fifth International Conference on Soft Computing for Problem Solving: SocProS 2015, Volume 1, Springer, 2016, pp. 333–343.
- [41] R. NALLAPATI, B. ZHOU, C. GULCEHRE, B. XIANG, ET AL., *Abstractive text summarization using sequence-to-sequence rnns and beyond*, arXiv preprint arXiv:1602.06023, (2016).
- [42] H. NIDA, K. MAHIRA, M. MUDASIR, M. MUDASIR AHMED, AND M. MOHSIN, *Automatic emotion classifier*, in Progress in Advanced Computing and Intelligent Engineering: Proceedings of ICACIE 2017, Volume 1, Springer, 2019, pp. 565–572.
- [43] R. PAULUS, C. XIONG, AND R. SOCHER, *A deep reinforced model for abstractive summarization*, in International Conference on Learning Representations, 2017.
- [44] A. M. RUSH, S. CHOPRA, AND J. WESTON, *A neural attention model for abstractive sentence summarization*, in Empirical Methods in Natural Language Processing (EMNLP), 2015, pp. 379–389.
- [45] A. SEE, P. J. LIU, AND C. D. MANNING, *Get to the point: Summarization with pointer-generator networks*, in Association for Computational Linguistics, 2017, pp. 1073–1083.
- [46] L. SONG, J. ZHANG, AND X. WAN, *Abstractive document summarization with a graph-based attentional neural model*, Information Processing & Management, 57 (2020), p. 102274.
- [47] A. VASWANI, N. SHAZEER, N. PARMAR, J. USZKOREIT, L. JONES, A. N. GOMEZ, AND I. POLOSUKHIN, *Attention is all you need*, in Advances in Neural Information Processing Systems, vol. 30, 2017.
- [48] E. VORHEES AND D. GRAFF, *AQUAINT-2 Information-retrieval Text: Research Collection*, Linguistic Data Consortium, 2008.
- [49] S. WEI, Y. HU, Z. CAI, AND K. ZHENG, *Multi-modal summarization for video-based learning content*, Information Processing & Management, 58 (2021), p. 102440.
- [50] S. YANG, J. YIM, J. KIM, AND H. V. SHIN, *Catchlive: real-time summarization of live streams with stream content and interaction data*, in Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, 2022, pp. 1–20.
- [51] J. ZHANG, T. WANG, AND X. WAN, *Pkusumsum: a java platform for multilingual document summarization*, in Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: System Demonstrations, 2016, pp. 287–291.
- [52] Y. ZHAO, X. MAO, AND Q. CUI, *Domain-specific extractive and abstractive summarization for academic papers*, Journal of Computer Science and Technology, 34 (2019), pp. 770–783.

*Edited by:* Mudasir Mohd

*Special issue on:* Scalable Computing in Online and Blended Learning Environments: Challenges and Solutions

*Received:* Aug 29, 2023

*Accepted:* Sep 21, 2023



## THE EFFECTS OF INTEGRATED FEEDBACK BASED ON AWE ON ENGLISH WRITING OF CHINESE EFL LEARNERS

MEI LIU\* AND CHANGZHONG SHAO†

**Abstract.** Through a two-semester experiment on the English writing of 64 Chinese EFL learners, this study examines the effects of three types web-based feedback (automatic feedback (AF), automatic feedback with teacher feedback (TF), automatic feedback with peer feedback (PF)) based on Pigai website. The results show that all the three modes of feedback can promote the writing of the English as Foreign Language (EFL) learners with different English proficiency ((high level:  $F = 2.672$ ,  $P = .132$ ; low level:  $F = .388$ ,  $P = .766$ ). The results also reveal that there is a significant difference in the high-level group between the automatic feedback + peer feedback group and the automatic feedback group ( $I - J = -6.636$ ,  $P = .000$ ) as well as between the automatic feedback + teacher feedback group and the automatic feedback group ( $I - J = -6.220$ ,  $P = .001$ ;  $I - J = -5.100$ ,  $P = .001$ ), which indicate that automatic feedback + manual feedback (PF+TF) can promote the improvement of high-level learners' English writing more than single AF. In the low-level group, between the AF + PF group and the AF group ( $I - J = -1.221$ ,  $P = .925$ ) and between the AF +TF group and the AF + PF group ( $I - J = 6.227$ ,  $P = .097$ ). There is no significant difference, but there is a significant difference between the AF group and the AF + TF group ( $I - J = -5.122$ ,  $P = .032$ ), indicating that AF + TF is of great help in improving the English writing of low-level English learners.

**Key words:** integrated feedback; automated writing evaluation (AWE); English writing; Pigai website

**1. Introduction.** Since the 1990s, feedback has been an important research topic in the field of second language acquisition, with the effectiveness of writing feedback being an important topic in second language writing feedback research [41, 42, 13, 24, 30]. In existing empirical studies, comparing the effectiveness of different feedback is a hot topic in writing feedback research [6, 7]. As Huisman et al. (2019) found, there was no significant difference in the impact of peer feedback (PF) and teacher feedback (TF) on students' writing performance. Thirakunkovit & Chamcharatsri [40] found that TF has a greater effect than PF. Double et al. (2020) [12] found that PF has a greater effect than TF. As to the research in China, Xing and Luo (2014) [54] found that PF and TF have a similar impact on students' overall writing level, and the effect of PF+TF is better than that of TF alone. Li and Sun (2019) [27] found that TF and mixed feedback have a significant effect, while PF has a relatively small effect. It can be seen that the effectiveness of feedback from different sources in writing is still controversial and requires further research in the EFL teaching context in China. Furthermore, in current English writing teaching in China, there are a lot of work for teachers to correct their learners' compositions. Although timely and effective TF is helpful to improve their learners' writing competence, it is often time-consuming and laborious. In order to solve this problem, Automated writing evaluation (AWE) system was born as needed. AWE can provide writing feedback on a large scale and in time, and to some extent, it can relieve the pressure of teachers' correction, so it has been rapidly popularized in the English as Foreign Language (EFL) teaching in China these years. Nowadays, some famous AWE systems including Pigai website, Criterion, MyAccess in China provide not only the total score and sub-score of the compositions, but personalized feedback from macro (such as essay center, structure, content organization, etc.) to micro (such as grammar, spelling, idiomatic usage, etc.) feedback. An in-depth comparison of the effects of automated feedback (AF), TF and PF on the EFL learners in China through empirical research is helpful for the active selection of the EFL writing teaching modes and methods in both China and other countries.

---

\*School of Education Science, Nanjing Normal University, Nanjing Jiangsu 210024, China; School of Education, Linyi University Linyi Shandong 276000, China

†School of Foreign Languages, Linyi University, Linyi Shandong 276000, China (Corresponding author, [shaochangzhong@lyu.edu.cn](mailto:shaochangzhong@lyu.edu.cn))

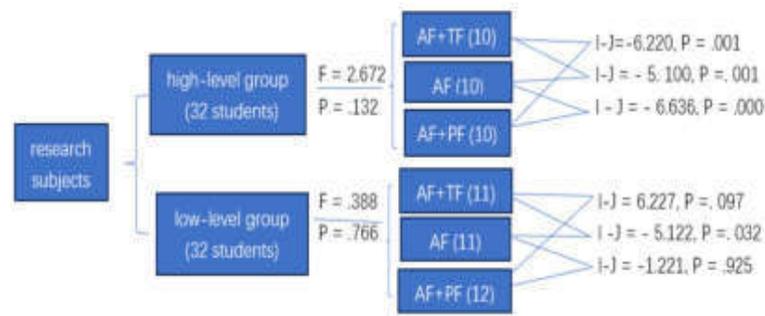


Fig. 1.1: Research Subjects

**2. Literature review.** Writing feedback is undoubtedly an important part of writing in both second language teaching and foreign language teaching around the world. How to evaluate the impact of feedback on language capacity in English as second language (ESL) and EFL writing, the past research has mainly focused on the impact of corrective feedback on writing accuracy or using CAF (Complexity, Accuracy and Fluency) as a framework to “panoramicly” measure and describe the accuracy, complexity and fluency of language in learners’ writing [3, 16, 29, 35, 45, 53]. Most of the research results found that writing feedback can promote learners to discover and revise problems, and get inspiration from feedback. It is an important tool for gradually developing self-expression capacity and a way for readers to respond and evaluate authors [47]. Feedback generally contains two basic elements: evaluation and correction. Evaluation refers to readers’ overall and general comments or ratings on the author’s article; revision refers to readers providing detailed introductions, explanations and guidance to the author, aiming to help the author identify and correct deficiencies [36]. The types of feedback in English writing can be distinguished from different perspectives: according to the form of feedback, it can be divided into oral feedback and written feedback; according to the source of feedback, it can be divided into teacher feedback, peer feedback and automatic feedback from online systems.

Writing feedback is an important part of the writing process. Appropriate feedback can help learners understand the strengths and weaknesses of their own writing and play a positive role in improving learners’ English writing [5]. In recent years, many researchers have paid more and more attention to the influence of writing feedback on improving English writing level, and different forms of feedback models have also been used in English writing teaching.

In tertiary English writing teaching, due to the limited time spent on writing training in the classroom, it is necessary to find a feedback method that allows learners to regularly practice English writing outside the classroom and receive immediate and effective feedback [2]. With the use of computer feedback and network feedback in English writing teaching, automatic writing evaluation system came into being. The automated writing evaluation (AWE) system based on computer technology has higher reliability and validity than manual evaluation [25, 50], and with other advantages like the immediacy of feedback (Attali 2004; Dikli 2006), attracting and promoting multiple revisions by learners [49], being able to be used for both formative and summative evaluations [39], so it has become a category that cannot be ignored in writing feedback research.

There are some representative writing automatic evaluation systems nowadays as Criterion, MyAccess, Writing Roadmap (WRM), and the most widely used composition automatic evaluation system in China is Juku Pigai Website (hereinafter referred to as Pigai website) developed by Beijing Ciwang (word network) Technology Co., Ltd. The research on the application of automatic evaluation system (AWE) in writing classrooms shows that the system can meet the needs of learners’ personalized evaluation and feedback, and help learners to correct language errors [34], and significantly improve composition scores and improve composition

quality [1, 17]. The use of automatic evaluation systems allows learners to continuously participate in the “feedback-practice-feedback” writing cycle, thereby continuously improving their writing skills [26].

Scholars hold different opinions on whether AWE can promote the development of learners’ written language. Some related studies have shown that the automatic scoring system can provide immediate and effective feedback to learners’ composition, so that the learners can make repeated revisions, improve their autonomous learning capacity, and reduce the workload of teachers [4, 38]. Some scholars have carried out practice and research on automatic composition evaluation system [49], and found that automatic feedback has a positive effect on the development of learners’ writing capacity because it can promote learners to revise multiple times, which proves that it can help improve learners’ English writing level used Criterion as a tool to provide automatic feedback on writing [46, 57], and found that AWE was beneficial to the improvement of learners’ language accuracy.

Nevertheless, it is impossible for machines to completely replace people, and the automatic review system has certain shortcomings. Studies proposed that the automatic revision of writings by machines deviates from the essence of writing itself [8, 33], which is a complex and highly interactive interpersonal communication behavior, and may mislead learners to focus only on syntactic form, while ignoring the development of content and ideas. On the one hand, “it can only judge the level from the language, not the content from the semantics, but it can’t identify the content errors of the composition, and some language errors can’t be identified” [22]. On the other hand, there are many studies on feedback from automated correction systems that are inconsistent. There is research studied the role of teachers and peers in writing evaluation, and the results showed that the learners who received teacher and peer feedback performed best [43]; the learners who received teacher feedback performed better than those who received peer feedback, while the learners who received peer feedback performed the worst. However, a study came to the opposite conclusion: the writing performance of the learners in the peer evaluation group has made greater progress, which is significantly better than that in the teacher feedback group [9]. With the deepening of the research, the researchers found that although the automatic evaluation system provides timely and fast feedback, it is inferior to the teacher feedback in terms of the quantity and accuracy of misidentification [10]. The system has played a positive role in correcting learners’ language errors, but the feedback information in terms of discourse content and coherence is vague, abstract, and programmed, and it cannot make specific revision suggestions for individualized problems like teachers do [7, 57].

At present, the AWE system represented by Pigai website has gradually been widely used in foreign language teaching in colleges and universities in China, especially in EFL teaching. However, there are few studies on the impact of automatic writing feedback on learners’ written language. A study used the Writing Roadmap 2.0 composition automatic evaluation tool developed in the United States to evaluate the impact of AWE on the development of Chinese college learners’ English writing capacity from 7 indicators of English composition [21], and made a positive conclusion to AWE. While [56], who also used Pigai website as a tool to study learners’ self-learning behavior in online writing, came to different conclusions. The former believes that the self-writing teaching mode based on Pigai website can effectively improve the overall level of learners’ English writing, but the latter finds that the number of learners’ self-correction is limited, the quality is not high, and the effect of composition optimization is not quite obvious.

In view of this, it has become one of the hot research issues in academia as how to combine the AWE system with teacher feedback to build a multiple feedback mechanism, to make up for the shortcomings of the AWE system by taking advantage of PF and TF, so as to provide learners with more comprehensive, objective and accurate feedback. A former study discussed the multi-dimensional feedback writing teaching pilot model based on the WRM automatic evaluation system [39]. The results show that this model can promote the development of learners’ writing capacity and has a positive impact on the teaching process. This research mainly focuses on the influence of multi-dimensional feedback teaching mode on learners’ writing capacity, but does not specifically discuss the influence of multi-feedback on learners’ writing modification. Moreover, the research object is high school learners, and the application and effect of other sample objects also need to be further tested. Huang’s works discussed the influence of multiple feedbacks such as feedback from Pigai website, peer feedback and teacher feedback on learners’ writing revision and writing behavior, and found that integrated feedback can improve learners’ writing behavior and motivate learners’ subjectivity to improve text quality [18, 19]. These two studies have conducted useful exploration and research on the influence of multiple

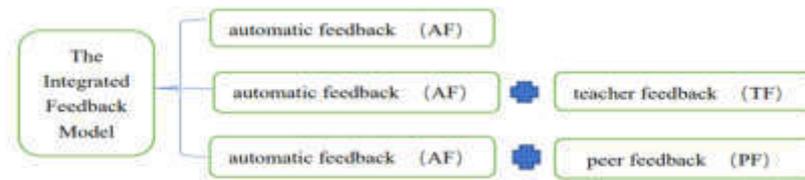


Fig. 3.1: The model of variables

feedback on the type and function of learners' writing revision and text quality, but did not conduct specific research on the impact of multiple feedback on the effect of learners' writing modification. Other studies have shown that, compared with the feedback provided by the teacher, the feedback provided by the computer is more effective in promoting language acquisition [47]. As a writing teaching and evaluation method, multiple feedback needs more theoretical and practical demonstrations, and more extensive sample experiments are needed to test its effects. Based on this, this study uses the feedback function of the automatic evaluation system of Pigai website, combines teacher feedback and student feedback, as well as the impact on the number, type and effect of revisions in learners' English writing, with a view to further enriching writing through this research. The content and scope of feedback research provide empirical evidence and a new feedback operation mode for English writing teaching.

### 3. Research design.

**3.1. Research questions.** This study attempts to answer the following questions:

1. Do the three feedback modes of AF, AF + TF and AF + PF have different effects on the EFL learners' English writing in China?
2. Is it related to the learners' English level (higher English proficiency and lower English proficiency)?

**3.2. Research method.** This study adopts a between-group design with a total of two independent variables: feedback mode (AF, AF + TF, AF + PF) and the subjects' English level (high and low level). The dependent variable is the final exam essay scores in semester 1 and semester 2.

**3.3. Research subjects.** The subjects of the study were two classes with 64 sophomore learners majoring in English at a comprehensive university in northern China, with 32 learners in each class, including 15 boys and 49 girls, with an average age of 20 years old. Before the experiment, the teacher assigned an English composition titled "Why Do You Learn English" to the learners in the two classes, which was required to be completed within 30 minutes and the number of words should not be less than 200 words. The composition is graded by two teachers at the same time. The grading standard is that each sub-item of length, structure, content and language expression accounts for 25 points, and the full score is 100 points. The average scores of the two classes were 73.60 (Class 1) and 68.63 (Class 2), and the scores of the two classes were significantly different ( $t = -7.332$ ,  $P = .000$ ). Therefore, the author set the learners of Class 1 (32 people) as the high-level group, and the learners of Class 2 (32 people) as the low-level group, and then randomly divided the learners of the high-level and low-level groups into six groups, respectively. Accept three different forms of feedback, namely automatic feedback (10 high level learners, 11 low level learners), automatic feedback + teacher feedback (10 high level learners, 11 low level learners), automatic feedback + peer feedback (10 high level learners, low level 12 learners).

**3.4. Selection of AWE system.** At present, there are many mature AWE systems in the world, but some are not designed for learners of second language acquisition or EFL learning, and it is difficult to provide feedback, textual and syntactic suggestions in line with the characteristics of the EFL learners in China. Therefore, this study wishes to adopts a self-developed AWE system by Chinese, which would be convenient for

its promotion in English teaching and learning in China, and it can also accumulate more empirical research experiences for the development of AWE system for the EFL learners in the national conditions of China.

After comparing the functions and convenience of several major AWE systems in China, we chose Pigai website, which is an online service system for automatic correction of English writings based on corpus and cloud computing. The analysis results of learners' writing can be generated instantly based on its corpus [22]. Like other AWE systems, Pigai website can provide individualized feedback on the total score, overall evaluation, vocabulary, grammar, style and other individual items of the writing. At the same time, it has its own innovations, such as the combination of automatic review of the support system and manual review by teachers, providing excellent writing sharing and some other functions.

**3.5. Research process.** In order to improve the evaluation capacity of learners' English writing and ensure that the peer feedback is really effective, the learners in the automatic feedback + peer feedback group were given peer feedback training before the experiment. It mainly trains learners how to give feedback on writing content to make up for the lack of feedback from the automatic review system. The feedback of the automatic rating system is mainly for the language form of the writing, including grammar, sentence pattern, vocabulary, punctuation, etc., and cannot give sufficient feedback on the structure, logic and coherence of the content. The content is the most important aspect to measure the quality of writing [51]. Only by paying full attention to the content, that is, paying attention to those elements with "substantiality" and "overall", can the writing capacity be improved [55].

In the experimental stage, the learners in two classes have to complete 10 writings in two semesters, 5 in each semester, and one composition every three weeks. The genre of these writing involves culture, life, morality, environmental protection, tourism and other hot social issues.

The teacher arranges the writing task on Pigai website and tells the composition number to the learners. After the learners in the automatic feedback group complete the composition, Pigai website will automatically correct it, and then the learners will modify it according to the automatic feedback until they are satisfied. After the learners in the AF + TF group complete the composition, they will also get the automatic feedback in Pigai website, and then the learners will modify it themselves until they are satisfied, and then the teacher will use the quick comment function of Pigai website to give feedback on the learners' composition, and then the learners will revise it again themselves; After the learners in the automatic feedback + peer feedback group complete the composition, and firstly the learners will revise the composition themselves according to the feedback by Pigai website until they are satisfied. Up to now, the teacher will randomly arrange mutual correction among learners on Pigai website, and then the learners will make corrections again themselves.

**3.6. Test score.** The research measurement objects of this study include: a pre-writing test and two post-writing tests (post-test 1 and post-test 2), in which the two post-tests are the final exam composition scores of the first and second semesters respectively.

Two teachers with extensive teaching experience will grade the learners' writing in the pre-test and the two post-tests. The scoring standard is that each sub-item of length, structure, content and language expression accounts for 25 points, and the full score is 100 points. When two teachers have a large difference in the scores of the same composition, the third teacher will give the score, and the average of the two similar scores will be taken at the end. If the difference is not large, the final score will be the average of the scores given by the two teachers. The researcher counted the reliability between the raters and found that the values were between 0.81 and 0.91, indicating that there was a high degree of consistency among the raters.

**4. Data analysis.** Table 4.1 shows the average scores and standard deviations of the writing scores of the high- and low-level groups under the three feedback methods of pre-test, post-test 1, and post-test 2.

The data in Table 4.1 shows that in the low-level group, the automatic feedback + teacher feedback group has the highest average scores in post-test 1 and post-test 2; in the high-level group, the automatic feedback + teacher feedback group and the automatic feedback + peer feedback group indicating that different feedback methods have different effects on learners' writing.

To further examine whether the three feedback methods have different effects on learners' writing, one-way ANOVA on the writing scores of post-test 1 and post-test 2 was performed respectively (see Table 4.2). The results showed that in post-test 1, there was no significant difference in the effects of the three feedback methods

Table 4.1: The writing scores of the high level and low level groups under the three feedback in the pre-test, post-test 1, and post-test 2 writing

Subject	Group	Number of learners	Pre-test		Post-test 1		Post-test 2	
			Average score	Standard deviation	Average score	Standard deviation	Average score	Standard deviation
Low level group	AF	11	75.53	3.54	77.21	4.55	79.52	4.82
	AF+TF	11	74.16	4.83	78.44	5.34	81.22	4.75
	AF+PF	12	76.87	4.41	78.63	5.46	81.74	5.58
	Total	34	75.52	4.26	78.09	5.12	80.82	5.05
High level group	AF	10	87.26	4.65	88.34	3.32	90.23	3.11
	AF+TF	10	88.38	4.86	89.77	3.34	90.73	2.67
	AF+PF	10	90.12	4.23	91.01	3.51	92.22	2.56
	Total	30	88.57	4.58	89.71	3.39	91.06	2.77

AF: Automated Feedback; TF: Teacher Feedback; PF: Peer Feedback

Table 4.2: One-way ANOVA results for post-test 1 and post-test 2

		Sum of squares		Degree of freedom		Mean square		F value		P value	
		High	Low	High	Low	High	Low	High	Low	High	Low
Post-test 1	Between groups	63.522	22.645	2	2	31.266	11.262	2.672	.388	.131	.766
	Within group	322.156	946.452	36	43	13.663	31.561				
	Total	385.678	969.097	38	45						
Post-test 2	Between groups	266.672	298.422	2	2	120.335	143.366	17.762	6.425	.000	.022
	Within group	233.362	928.322	36	43	8.336	32.163				
	Total	500.034	1226.744	38	45						

\*  $P < .05$  means there is a significant difference

on the writing of high-level and low-level learners (high level:  $F = 2.672$ ,  $P = .132$ ; low level:  $F = .388$ ,  $P = .766$ ); and in the post-test 2, the effects of the three feedback methods on the writing of high-level and low-level learners were significantly different (high level:  $F = 17.762$ ,  $P = .000$ ; low level:  $F = 6.425$ ,  $P = .022$ ), indicating that different feedback methods have different effects on learners' writing.

We performed a post-hoc test of the effect of the three types of feedback on writing in post-test 2 with the Schéffe test (see Table 4.3). The results showed that, there is a significant difference in the high-level group between the automatic feedback + peer feedback group and the automatic feedback group ( $I - J = -6.636$ ,  $P = .000$ ) as well as between the automatic feedback + teacher feedback group and the automatic feedback group ( $I - J = -6.220$ ,  $P = .001$ ;  $I - J = -5.100$ ,  $P = .001$ ), which indicate that automatic feedback + manual feedback (PF+TF) can promote the improvement of high-level learners' English writing more than single automatic feedback. In the low-level group, between the automatic feedback + peer feedback group and the automatic feedback group ( $I - J = -1.221$ ,  $P = .925$ ) and between the automatic feedback + teacher feedback group and the automatic feedback + peer feedback group ( $I - J = 6.227$ ,  $P = .097$ ) There is no significant difference, but there is a significant difference between the automatic feedback group and the automatic feedback + teacher feedback group ( $I - J = -5.122$ ,  $P = .032$ ), indicating that AF + TF is of great help in improving the English writing of low-level English learners.

In order to verify whether the three types of feedback have an impact on the writing effect, the author conducted a paired sample t-test on the English writing scores in the pre-test, post-test 1 and post-test 2. It can be seen from Table 4.4 that the last test of the three types of feedback is significantly different from the

Table 4.3: The influence of three different feedback methods on the English writing capacity of the high- and low-level groups: results of Scheffé test analysis

Subject	dependent variable	independent variable		Mean difference (I-J)	P value
		(I) Feedback	(J) Feedback		
Low level group	Post-test 2	AF	AF+TF	-5.122(*)	.032
		AF	AF+PF	-1.221	.925
		AF+TF	AF+PF	6.227	.097
High level group	Post-test 2	AF	AF+TF	-6.220(*)	.001
		AF	AF+PF	-6.636(*)	.000
		AF+TF	AF+PF	-.766	.935

AF: Automated Feedback; TF: Teacher Feedback; PF: Peer Feedback

\* $P < .05$  means there is a significant difference

Table 4.4: Paired sample t-test between pre-test, post-test 1, post-test 2

Subject	Group	Test pairing	t-value	p-value
Low level group	AF	pretest-posttest 1	-3.954	.002
		pretest-posttest 2	-6.883	.000
		Posttest 1-posttest 2	-9.788	.000
	AF+TF	pretest-posttest 1	-6.667	.000
		pretest-posttest 2	-12.465	.000
		Posttest 1-posttest 2	-7.232	.000
	AF+PF	pretest-posttest 1	-6.645	.000
		pretest-posttest 2	-10.112	.000
		Posttest 1-posttest 2	-8.267	.000
High level group	AF	pretest-posttest 1	-7.757	.001
		pretest-posttest 2	-3.326	.000
		Posttest 1-posttest 2	-6.563	.017
	AF+TF	pretest-posttest 1	-6.672	.015
		pretest-posttest 2	-9.553	.000
		Posttest 1-posttest 2	-4.425	.012
	AF+PF	pretest-posttest 1	-5.344	.032
		pretest-posttest 2	-11.332	.000
		Posttest 1-posttest 2	-6.733	.004

previous test, indicating that the three types of feedback can promote the English writing level of both high- and low-level learners.

## 5. Results and discussion.

**5.1. Results.** The experimental results show that, in post-test 2, there are significant differences in the effects of the three modes of feedback on the English writing of high-level learners and low-level learners. For high-level learners, there are significant differences between AF + PF group and the AF group, as well as between AF + TF group and AF group, indicating that AF + manual feedback mode is better than single AF, and manual feedback makes up for the lack of AF. However, there is no significant difference between AF + PF group and AF + TF group, which may be because the learners in the high-level group have stronger evaluation capacity, although it may not be as good as TF, it can also prompt peers to improve their English writing level. At the same time, learners can also find problems in their own writing while evaluating other learners' writing, so as to improve their own writing. Rollinson (2005) believes that learners can learn to review and revise their

own articles through peer review (Rollinson 2005). Both Tsui and Lundstroms found that reviewing peers' compositions was more beneficial to improving writing skills than receiving PF (Lundstroms and Baker 2009; Tsui and Ng 2000), because reviewers could learn some writing skills from others' compositions and apply them to their own writing to improve writing skills.

For low-level subjects, there was no significant difference between AF + PF group and AF group. This may be because the learners with a lower level cannot find problems with each other, and even if they find problems with the help of their peers, they are not able to put forward specific suggestions for revision, so these learners basically fail to meet the requirements of PF. In addition, it may also be that learners lack of confidence in peers, believing that they cannot be a qualified reviewer, and therefore be skeptical of PF (Cai 2011; Muncie 2000; Tsui and Ng 2000), and thus slower to improve writing skills (Gong 2007; Wang 2004). There is a significant difference between AF group and AF + TF group, indicating that TF is of great help to the improvement of English writing skills of the learners with lower levels. However, in post-test 1, there was no significant difference in the effects of the three types of feedback on the English writing of the high-level and the low-level learners, probably because the experiment time was not long enough to find the difference in the effect between different feedback modes.

**5.2. Discussion.** This study also found that there were significant differences between the last test of the three modes of feedback and the previous test, and the three modes of feedback had a positive effect on the improvement of English writing capacity of both high-level and low-level learners in China. The automatic evaluation system can greatly reduce the workload of teachers, and it can be used in a reasonable combination with teacher feedback to give full play to its potential. It can reduce the burden on teachers, promote learners' writing practice, and improve their English writing faster. In fact, compared with face-to-face feedback, online writing feedback has less psychological pressure and less anxiety, so learners are more willing to put forward real ideas and opinions (Guardado et al. 2007; Jiang 2005).

This paper conducted a specific study on the impact of diverse feedback on the effectiveness of student essay revision, while Huang and Zhang (2014) and Huang and He (2018) only explored the impact of diverse feedback on the types and functions of student essay revision, as well as text quality. Although some study has shown that feedback provided through computers is more effective in promoting language acquisition compared to feedback provided by teachers (Wang Lina et al., 2018: 44), and this study combines sample experiments to test its effectiveness, providing theoretical and practical evidence for multiple feedback. Furthermore, this study also provides empirical evidence for further enriching the content and scope of writing feedback study by using the method of natural classroom data collection, which also provides a new feedback operation mode for English writing teaching and evaluation methods.

**6. Conclusions.** Multiple feedback based on automatic evaluation systems emphasizes the supervision and regulation of students' writing process. By adopting a feedback model that combines AF, TF, and PF, the learning space can be expanded, and various feedback can be organically combined and mutually supplemented, thereby stimulating students' initiative and enthusiasm for writing revision. Moreover, manual-machine collaborative review can incorporate interpersonal interaction into writing feedback, overcoming the limitation of automatic writing evaluation systems that only focus on cognitive processing, thus effectively achieving the purpose of writing communication in both form and meaning.

This study still has certain limitations: First, the subjects are just the EFL learners in a local university in China, and the number of research subjects is small. The high-level and low-level groups in the experiment are only relative, and further study on the effects of the EFL learners' English writing feedback in China with other levels of subjects are expecting. Second, this study explores the overall quality of the EFL learners' English writing, but it does not explore specific dimensions of English writing, such as accuracy, fluency, language diversity, and syntactic complexity. More collaborating study on these dimensions are expected in the future.

**Acknowledgements.** This article is the periodical research of the project 'The influencing factors and effective paths of blended learning development in universities from the ecological perspective' (Code: J18RA149) supported by the Shandong Education Department.

## REFERENCES

- [1] Attali, Y. Exploring the feedback and revision features of Criterion. *Journal Of Second Language Writing*. **14** pp. 191-205 (2004)
- [2] Bitchener, J. & Knoch, U. The value of a focused approach to written corrective feedback. *English Language Teachers Journal*. **3** pp. 204-211 (2009)
- [3] Bitchener, J. & Young, S. The effect of different types of corrective feedback on ESL student writing. *Journal Of Second Language Writing*. **14** pp. 191-205 (2005)
- [4] Black, P. & Wiliam, D. Assessment and classroom learning. *Assessment In Education*. **1** pp. 7-74 (1998)
- [5] Cai, J. A comparative study of online peer feedback and teacher feedback on Chinese college students. *English Writing', Foreign Language World*. **2** pp. 65-72 (2011)
- [6] Chang, C. & Chang, H. Retrospect and prospect of EFL writing research in China. *Technology Enhanced Foreign Language Education*. **193** pp. 61-67 (2020)
- [7] Chen, C. & Cheng, W. Beyond the design of automated writing evaluation: pedagogical practices and perceived learning effectiveness in EFL writing classes. *Language Learning & Technology*. **2** pp. 94-112 (2008)
- [8] Chevillat, J. Automated scoring technologies and the rising influence of error. *The English Journal*. **4** pp. 47-52 (2004)
- [9] Deng, L. & Cen, Y. A study on the efficiency of peer review feedback mechanism of L2 writing proficiency development. *Foreign Language Education*. **31** pp. 59-63 (2010)
- [10] Dikli, S. An overview of automated scoring of essays. *Journal Of Technology, Learning, And Assessment*. **5** pp. 1-35 (2006)
- [11] Dikli, S. & Bleyle, S. Automated essay scoring feedback for second language writers: how does it compare to instructor feedback?. *Assessing Writing*. **22** pp. 1-17 (2014)
- [12] Double, K., Mcgrane, J. & Hopfenbeck, T. The impact of peer assessment on academic performance: A Meta- analysis of control group studies. *Educational Psychology Review*. **32** pp. 481-509 (2020)
- [13] Ferris, D. Written corrective feedback in second language acquisition and writing studies. *Language Teaching*. **45** pp. 446-459 (2012)
- [14] Gong, X. Teaching English writing: optimized peer feedback. *Foreign Language Teaching Abroad*. **3** pp. 47-51 (2007)
- [15] Guardado, M. & Shi, L. ESL students experiences of online peer feedback. *Computers And Composition*. **24** pp. 443-457 (2007)
- [16] Hartshorn, K., Evans, J., N. W., M., P. F., S. & R. R., S. D. and Anderson, N. J. 'Effects Of Dynamic Corrective Feedback On ESL Writing Accuracy', *TESOL Quarterly*. **44** (2010)
- [17] Hu, X. The effects of online feedback to English writing of college students. *Technology Enhanced Foreign Language Education*. **3** pp. 45-49 (2015)
- [18] Huang, J. & He, H. A study on the influence of human-computer feedback on students. *Writing Behavior', Technology Enhanced Foreign Language Education*. **1** pp. 19-24 (2018)
- [19] Huang, J. & Zhang, W. A study of multiple feedback to the correction of college Students. *English Writing', Foreign Languages In China*. **1** pp. 51-56 (2014)
- [20] Huisman, B., Saab, N., Broek, P. & Driel, J. The impact of formative peer feedback on higher education students. *Academic Writing: A Meta- Analysis', Assessment & Evaluation In Higher Education*. **44** pp. 863-880 (2019)
- [21] Jiang, X., Cai, J. & Tang, J. To explore the effects of AWE to the English writing competence of college students. *Shandong Foreign Language Teaching Journal*. **6** pp. 36-43 (2011)
- [22] Jiang, Y. & Ma, W. Achievements and challenges of Chinese English writing teaching intelligent tutor system: taking Pigai website as an example. *Electronic Education Research*. **7** pp. 76-81 (2013)
- [23] Jiang, Y. The role of online peer review in learners. *Writing Ability Development', Foreign Language Teaching And Research*. **37** pp. 226-230 (2005)
- [24] Kang, E. & Han, Z. The efficacy of written corrective feedback in improving L2 written accuracy: A meta-analysis. *Modern Language Journal*. **99** pp. 1-18 (2015)
- [25] Keith, T. Validity of automated essay scoring systems. *Automated Essay Scoring: A Cross-disciplinary Perspective*. (2003)
- [26] Kellogg, R., Whiteford, A. & Quinlan, T. Does automated feedback help students learn to write. *Journal Of Educational Computing Research*. **2** pp. 173-196 (2010)
- [27] Li, F. & Sun, Y. A Meta- analysis of the effectiveness of written corrective feedback based on domestic literature. *Foreign Language Research*. **210** pp. 55-62 (2019)
- [28] Lundstroms, K. & Baker, W. To give is better than to receive: the benefits of peer review to the reviewer. *S Own Writing', Journal Of Second Language Writing*. **18** pp. 30-43 (2009)
- [29] Masatoshi, S. & Roy, L. Peer interaction and corrective feedback for accuracy and fluency development. *Studies In Second Language Acquisition*. **34** pp. 591-626 (2012)
- [30] Mohebbi, H. 25 years on, the written error correction debate continues: An interview with John Truscott. *Asian-Pacific Journal Of Second And Foreign Language Education*. **6** pp. 1-8 (2021)
- [31] Muncie, J. Using written teacher feedback in EFL composition classes. *ELT Journal*. **54** pp. 47-53 (2000)
- [32] Nguyen, N., Xiong, W. & Litman, D. Iterative design and classroom evaluation of automated formative feedback for improving peer feedback localization. *International Journal Of Artificial Intelligence Education*. **27** pp. 582-622 (2017)
- [33] Perin, D. & Lauterbach, M. Assessing text-based writing of low-skilled college students. *International Journal Of Artificial Education*. **28** pp. 56-78 (2018)
- [34] Petchprasert, A. Utilizing an automated tool analysis to evaluate EFL students' writing performances. *Asian-Pacific Journal Of Second And Foreign Language Education*. **6** pp. 1-16 (2021)
- [35] Plakans, L., Gebril, A. & Bilki, Z. Shaping a score: Complexity, accuracy, and fluency in integrated writing performances.

- Language Testing*. **50** pp. 1-19 (2016)
- [36] Rahimil, Z., Litman, D., Correnti, R., Wang, E. & Matsumura, L. Assessing students. *Use Of Evidence And Organization In Response-to-text Writing: Using Natural Language Processing For Rubric-based Automated Scoring*, *International Journal Of Artificial Intelligence In Education*. **27** pp. 694-728 (2017)
- [37] Rollinson, P. Using peer feedback in the ESL writing class. *ELT Journal*. **59** pp. 23-30 (2005)
- [38] Sowmya, V. Automated assessment of non-native learner essays: investigating the role of linguistic features. *International Journal Of Artificial Education*. **28** pp. 79-105 (2018)
- [39] Tang, J. How to integrate an automated writing assessment tool in the EFL classroom. *Foreign Language Learning Theory And Practice*. **1** pp. 49-57 (2014)
- [40] Thirakunkovit, S. & Chamcharatsri, B. A meta-analysis of effectiveness of teacher and peer feedback: Implications for writing instructions and research. *Asian EFL Journal*. **21** pp. 140-170 (2019)
- [41] Truscott, J. The case against grammar correction in L2 writing classes. *Language Learning*. **46** pp. 327-369 (1996)
- [42] Truscott, J. The effect of error correction on learners. *Ability To Write Accurately*, *Journal Of Second Language Writing*. **16** pp. 255-272 (2007)
- [43] Tsui, A. & Ng., M. . *Do Secondary L*. **2** (2000)
- [44] Tsui, A. & Ng, M. Do secondary L2 writers benefit from peer comments?. *Journal Of Second Language Writing*. **2** pp. 147-170 (2000)
- [45] Ullmann, T. Automated analysis of reflection in writing: validating machine learning approaches. *International Journal Of Artificial Intelligence In Education*. **29** pp. 217-257 (2019)
- [46] Vajjala, S. Automated assessment of non-native learner essays: investigating the role of linguistic features. *International Journal Of J Artificial Intelligence Education*. **28** pp. 79-105 (2018)
- [47] Wang, L., Mo, L., Wu, Y. & Cappelletti, B. A meta-analysis study on the effectiveness of written corrective feedback. *Journal Of Xi'an International Studies University*. **26** pp. 40-45 (2018)
- [48] Wang, X. Can students master mutual modification skills?. *Foreign Language Teaching Abroad*. **1** pp. 54-56 (2004)
- [49] Warschauer, M. & Ware, P. Automated writing evaluation: defining the classroom research agenda. *Language Teaching Research*. **10** pp. 157-180 (2006)
- [50] Weigle, S. Validation of automated scores of TOEFL iBT tasks against non-test indicators of writing ability. *Language Testing*. **27** pp. 335-353 (2010)
- [51] Wen, Q. 'Measuring the validity of the construct "Essay Content" by using structure equation', *Foreign Languages Research*. (2007)
- [52] Wilson, J. & Czik, A. Automated essay evaluation software in English language arts classrooms: Effects on teacher feedback, student motivation, and writing quality. *Computers And Education*. **100** pp. 94-109 (2016)
- [53] Wolfe-Quintero, K. & Inagaki, S. Second language development in writing: measures of fluency, accuracy & complexit. *Manoa: University Of Hawaii, ISBN: 0. 8248* (1998)
- [54] Xing, J. & Luo, S. A comparative study of the effects of different forms of corrective feedback on Chinese college students. *English Writing: A Meta-analysis Based On Domestic Studies*, *Foreign Language Education In China*. **7** pp. 78-86 (2014)
- [55] Yang, M., Badger, R. & Yu, Z. A comparative study of peer and teacher feedback in a Chinese EFL writing class. *Journal Of Second Language Writing*. **3** pp. 179-200 (2006)
- [56] Yang, X. & Dai, Y. Study on the practice of college English autonomous writing teaching mode based on Pigai website. *Technology Enhanced Foreign Language Education*. **3** pp. 17-23 (2015)
- [57] Zhang, L. & Sheng, Y. A case study on the feedback effect of automatic composition review system. *Technology Enhanced Foreign Language Education*. **3** pp. 45-49 (2015)

*Edited by:* Mudasir Mohd

*Special issue on:* Scalable Computing in Online and Blended Learning Environments: Challenges and Solutions

*Received:* Sep 19, 2023

*Accepted:* Nov 17, 2023



## AN ELIXIR FOR BLOCKCHAIN SCALABILITY WITH CHANNEL BASED CLUSTERED SHARDING

V. VINOTH KUMAR\*, U. PADMAVATHI † C. PRASANNA RANJITH ‡ BALAJI J § AND C.N.S. VINOTH KUMAR ¶

**Abstract.** Blockchain refers to distributed ledger technology which stores records without the help of a central authority. Born with bitcoin, this brainstorming technology finds its applications in healthcare, land registry, education, pharmaceutical industry, digital records, manufacturing companies and so on. The properties of blockchain such as immutability, distributed nature, tamper-resistant made it a disruptive technology in many applications. The highlighting feature of this pioneering technology is the distributed storage of ledger on all the nodes of the network. This helps to achieve decentralization without the trust for third party. The transactions are proposed, executed, validated and are then added as blocks to the blockchain. The problems with all the blockchain framework is scalability with respect to storage space and throughput. Scalability is the most significant factor to be considered in this big data era. This article proposes a solution called Channel Based Clustered Sharding (CBCS) approach for Hyperledger fabric blockchain framework. In this work, a lookup table is maintained which helps in forwarding the transactions to the clustered shards for validation. The CBCS approach helps in parallel transaction processing which in turn improves scalability and throughput of the system. The performance of the proposed work is measured with the help of Hyperledger caliper, a benchmarking tool for the performance analysis of Hyperledger fabric. The results show that the performance of the proposed system is increased from 3000 tps to 30,000 tps.

**Key words:** Blockchain, Sharding, Scalability, cluster, channel, Hyperledger fabric, Hyperledger caliper

### 1. Introduction.

**1.1. Blockchain.** Blockchain is a distributed ledger technology that stores records in a chained manner without the need for trusted central authority [1]. Blockchain, the underlying concept of Bitcoin gained its importance in the world of cryptocurrencies. Satoshi Nakamoto, the founder of bitcoin used Proof of Work consensus mechanism to make consensus among the participants [2]. The proof of work consensus involves high transaction fees and more computation time. Since the ledger must be stored on all the participants in the network, it takes more time for processing the transactions and adding the blocks to the blockchain [3, 4, 5]. This results in scalability issue in bitcoin blockchain which can process only 7 transactions per second. Several protocols have been proposed to solve this issue [6, 7, 8]. With the Proof of Stake consensus, Ethereum can process 27 transactions per second. The modular architecture and execute-order-validate mechanism of Hyperledger fabric framework makes it possible to transact 3000 transactions per second which is very low when comparing to visa which makes 7000 transactions per second [9, 10]. It is found that the root cause of scalability issues in blockchain is the way the ledger gets stored on the nodes and the amount of time taken by all the nodes for validating the transactions. If every node in the network validates the transaction, then it would result in more consumption of time, which in turn reduces the parallel transaction processing and scalability. This work proposes channel based clustered sharding method to improve the scalability factor of Hyperledger fabric framework. Hyperledger fabric, one of the key projects of IBM differ from other blockchain network in the way it is designed. It has modular architecture which allows the user to plug-and-play with its

---

\*School of Computer Science Engineering and Information Systems, Vellore Institute of Technology, Vellore, Tamil Nadu, India; School of Computer Science, Taylor's university, Malaysia ([drvinothkumar03@gmail.com](mailto:drvinothkumar03@gmail.com))

†Department of Computer Science & Engineering, School of Engineering, Shiv Nadar University Chennai, India ([udayarajepadma@gmail.com](mailto:udayarajepadma@gmail.com))

‡Department of Information Technology, University of Technology and Applied Sciences-Shinas, Oman ([prasanna.christodoss@shct.edu.om](mailto:prasanna.christodoss@shct.edu.om))

§Department of Computer Science and Engineering, Mother Theresa Institute of Engineering Technology, India ([jbbala@gmail.com](mailto:jbbala@gmail.com))

¶ Department of Networking and Communications, College of Engineering and Technology (CET), SRM Institute of Science and Technology, Kattankulathur Chennai, India. ([vinothks1@srmist.edu.in](mailto:vinothks1@srmist.edu.in)).

components and it uses execute-order-validate mechanism which helps to improve transaction throughput of the system [11]. It makes use of practical byzantine fault tolerant consensus mechanism which provides security to the system. The channel concept in Hyperledger Fabric play a significant role in providing privacy among the users of the system. In this work, Hyperledger fabric framework has been chosen to improve the scalability. The objective of this work is to increase parallel transaction processing and reduce transaction latency. Various solutions has been proposed by various researchers to improve the scalability of the system [12, 13, 14]. This article proposes a database mechanism called sharding which is used to cluster the nodes into shards in each channel. The blockchain Sharding mechanism profoundly improves validation time which in turn improves parallel transaction processing and thus scalability.

**1.2. Sharding.** Sharding is a database mechanism that helps to split and store the data in different resources. It helps to manage data faster and in an efficient manner. The database is divided into number of smaller pieces and each piece of data is stored on different servers and that server is responsible for storing and managing the data. A shard key is used to track which data has been stored on which server. Sharding concept was introduced in blockchain in the year 2017. Luu et al [15] in his paper published the use of sharding in blockchain in which each block in the blockchain is divided into sub blockchain and each sub-blockchain can store several collation packaged with transaction data [16]. These collation constitute the block in the main chain. This work reduces additional network confirmation which in turn increased the trading capacity of block by 100 times. Elastico and Zilliqa use sharding [17] and they prove that the probability that the attacker can control sharding is very negligible. In case of Ethereum, it is in the process of incorporating sharding mechanism in its network. It works in layer 2 for dividing the processing of large chunks of data across the network thereby reducing the network congestion and improving transaction throughput. In this work, sharding is proposed to validate the transactions based on the channels. For each channel, separate clusters of shards are presented which helps in validating the transaction data providing confidentiality and security. It also increases the parallel transaction processing of the system by reducing the overhead involved in validating the data on all nodes in each channel. The main contributions of this work are as follows:

1. To design a channel based clustered sharding approach for hyperledger fabric blockchain framework.
2. To analyse the performance of the system using hyperledger caliper benchmark tool in terms of success rate, throughput and latency.

*Paper Organization:* Section 2 discusses the background and related work. Section 3 proposes the CBCS approach for solving scalability problem. Section 4 evaluates the proposed work and presents the results. Section 5 concludes the paper and discusses the future work.

**2. Background and Related Works.** This work reviews two different kinds of solution to blockchain scalability issues. One is on-chain solutions and the other is off-chain solutions. On-chain solutions concentrate on block structure, algorithms for consensus, main-chain structure. Off-chain focuses on reducing the burden of on-chain transactions by executing some complex transactions off-chain.

**2.1. On-chain Solutions.** SEGWIT [18] proposed the segregated witness based solution for blockchain scalability issue. It redesigns the structure of the transaction by computing the transaction identifiers without including the count of signatures. Segwit acts as a soft fork and helps to create bigger block sizes which in turn increase the scalability of the bitcoin protocol. Han et al proposed the Consensus unit in which different nodes are organized as one unit and the data are stored in one node of a unit. This helps to reduce the duplication of data in all the nodes in the network. The storage space is greatly reduced, and the transaction throughput is increased [19]. In JIGSAW approach [20], each node in the network stores only transactions that are of more interest. Every node maintains a local storage in which all the relevant data transactions are stored locally without the intervention of other nodes in the network. This system verifies the validity of the transaction with the help of merkle branch information provided by the transaction proposer. This approach helps to reduce the storage space cost to 1.03% of the original cost of blockchain bitcoin system. Lewenberg in his work proposes a DAG based structure of blockchain that helps to improve the transaction rate of the network [21]. The DAG structure is created by referring multiple predecessors and accepting forgiving transactions. It increases the transaction volume of the network by propagating longer blocks in the network. This system avoids highly connected miners because of which the block creation time is greatly reduced and concludes that the scalability



Fig. 3.1: Proposed System Architecture

of the system is greatly increased at the rate of security. [22] proposed a protocol called Phantom, based on Directed Acyclic Graph that supports faster and longer block generation. It overcome the security-scalability tradeoff using a greedy algorithm that helps to distinguish the blocks generated by honest nodes and blocks generated by non-cooperating nodes without following the mining protocol. This work guarantees liveness and fast confirmation times by the use of blockDAG.

**2.2. Off-chain solutions.** Poon J et al proposed a network of micropayment channels that enables bitcoin scalability. These micropayment channels are made off the chain and makes use of scripting opcodes that makes the transactions risk free and more scalable[23]. This work concludes that the bitcoin blockchain can scale upto millions of users without custodial risk or centralization. An off-chain blockchain transaction mechanism that creates long-lived channels for making arbitrary number of transactions between users in the network is presented in [24]. In this work, the payments are made between users without any confirmation delay. It guarantees end-to-end security between transactions in the bitcoin network. Sprites, a modular protocol approach that allows distrustful parties to make payments between them with reduced collateral cost is proposed by Miller et al. It makes use of generic state channel abstraction that helps to improve payment channel constructions and supports partial payments and withdrawal. Sprites ensures transaction scalability along with security in bitcoin network [25, 26]. HF-Audit, a decentralized data integrity auditing scheme based on Hyperledger Fabric is reported in [27]. In this work, bilinear pairing and commitments are made use of for improving the scalability and security of the Hyperledger fabric transactions at reduced cost. The security and scalability of HF-Audit is proved using Third-Party Audit selection algorithms based on complete and incomplete information. An adaptive framework that redesigns the consensus protocol of Hyperledger Fabric blockchain framework for improving the transaction throughput is proposed by Honar Pajooh. This framework checks the performance of Hyperledger fabric blockchain under various transaction workloads and different chaincode parameters. The impact on transaction latency and transaction throughput are studied and found that there is a significant improvement in the scalability of the system [28].

**3. Proposed System Architecture.** The proposed system architecture is depicted in the Fig. 3.1. It consists of peer nodes, endorser nodes, orderer nodes, channel 1, channel 2, certificate authority, channel configuration policy, shards  $C_1S_1$ ,  $C_1S_2$ ,  $C_2S_1$ , a lookup table and a ledger. The certificate authority in the Hyperledger fabric network is responsible for generating the certificates for the user. Using this certificate, the user can make transaction proposals and responses. The proposals from the user are signed with this certificate and are then transferred to the network. Along with this, in the proposed system the certificate authority is responsible for maintaining a look up table which helps the peer nodes to join the clustered shards and validate the transactions. The peer nodes are the nodes which is responsible for hosting the ledger data and the chaincode. The peer nodes are connected to the channels and each peer node stores the ledger data corresponding to the channels to which it is attached. The endorser nodes are the peer nodes which are responsible for endorsing the transactions proposed by the client applications. Only after every transaction gets endorsement according to the endorsement policies, the transactions be added to the blockchain. Orderer nodes are the special nodes

in the Hyperledger fabric framework which helps in ordering the transactions into blocks. In our proposed system we assume that there are two channels. With the help of channels, the Hyperledger fabric framework provides privacy and confidentiality in transaction communications. The transactions between the participants of a channel are not visible to participants in the other channel. Each channel maintains a ledger which contains only the transactions happened between the participants of that channel. The channels are governed by the channel confirmation policies.

Ledger is the place where the transactions are stored. Ledger comprises blockchain and the world state. World state represents the current values of the data. World state contains the key value pair which stores the latest value for each key and also it contains the version number. The version number is used by the Hyperledger fabric internally and it changes after each update made on the world state value. Blockchain contain the transaction logs which resulted in the current value of the data. It stores the history of transactions and the blocks are linked to each other with the help of hash values.

**3.1. Channel Based Clustered Sharding.** The channel based clustered sharding approach has been designed for drug supply chain applications. The participants involved in this application are the manufacturer, retailer, wholesale, pharmacist and the consumer. The manufacturer purchases the raw material and produces the drug which is then sold to the retailer. The retailer purchases the drug and sell it to the wholesale. The wholesaler in turn sells the drugs to the pharmacist based on their request. The pharmacist can also purchase drugs from the manufacturer directly. This transaction between the manufacturer and the pharmacist must be kept private without the knowledge of other participants involved in the network. This is possible with the help of channels in the proposed system. The consumer purchases the drug from the pharmacist. The proposed architecture involves two channel. Each channel is responsible for maintaining its own ledger. Channel 1 is dedicated for the transactions that happen between the manufacturer, retailer, wholesaler and the pharmacist. Channel 2 is dedicated for the transactions that happen between the manufacturer and the pharmacist.

**3.2. Shard Formation and Transaction Validation.** Formation of shard plays a vital role in the proposed system. Each shard consists of a cluster of nodes in the network. The proposed CBCS consists of three shards namely  $C1S1$ ,  $C1S2$  and  $C2S1$ . When the nodes join the network, the nodes will be randomly assigned to one of the shards in each channel. With the increase in the number of peers, the number of shards will also increase. The shards are grouped in such a way that, they will validate the transactions only from certain nodes. In this work, shard  $C1S1$  will validate the transactions from the manufacturer and the retailer nodes in channel 1. Shard  $C1S2$  will validate the transaction from the wholesaler and the pharmacist node in channel 1. The shard  $C2S1$  validate the transactions from the manufacturer and the pharmacist nodes in channel 2. This helps in parallel transaction validation which in turn increases the transaction throughput per second. Once the transactions are validated, they are then ordered and stored on to the blockchain.

**3.3. Transaction Flow in the proposed system.** The certificate authority is responsible for providing certificates to the nodes when they join the network. Using the certificate, the participants make transaction proposal in the system. This certificate also helps to verify the authenticity of the participants in the network. The certificate authority is also responsible for maintaining the look up table in which information about different shards and which nodes are present in which clustered shard are maintained.

When the retailer in channel 1 proposes a transaction tx, it is endorsed and the responses are sent back to the retailer. The transaction proposal and the response received are packed as a transaction and are sent to the orderer nodes. The orderer nodes arrange the transactions and propagate it to the nodes in shard  $C_1S_1$ . These nodes are responsible for validating the transaction and update the current status of the ledger. It is illustrated in Fig. 3.2. If any other transactions are received from the wholesaler in Channel 1 at the same time, then these transactions are validated by the nodes in shard  $C_1S_2$ . This helps to increase the parallel transaction processing of the system. The transactions generated in the channel 2 are validated by the nodes of the shard  $C_2S_1$  and the ledger is then updated. If the transaction generated is a query transaction, then these transactions are not sent to the ordering node to be added to the blocks and for validation. This method helps in reducing the transaction processing time and improving the parallel processing of transactions. It also improves the scalability of the system to a larger extent.

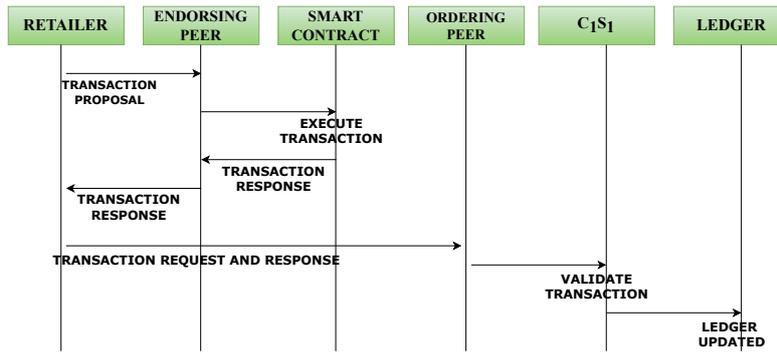


Fig. 3.2: Transaction Flow in the Proposed System

Transaction	Success Rate
1000	100
2000	100
3000	98
5000	96
10000	82
20000	75
30000	60
40000	54
50000	35

Table 4.1: Impact on Success Rate

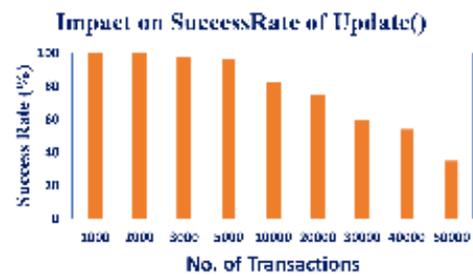


Figure 4.1: Impact on Success Rate of Update()

**4. Result.** The proposed system is tested with the help of a benchmarking tool called Hyperledger caliper. It supports Hyperledger fabric, Hyperledger sawtooth, Hyperledger besu and Ethereum blockchain framework. Hyperledger caliper produces results based on success rate, transaction throughput and transaction latency. The term success rate refers to number of successful transactions per test cycle.

The term transaction throughput is defined as the number of transactions that flow through the system per second. The term transaction latency is defined as the time interval between the transaction being completed and the response being available to the user application which created the transaction.

**4.1. Setting up Hyperledger Caliper.** To test the performance of the proposed work, certain parameters need to be set up in Hyperledger Caliper. The number of transactions, Rate control, Number of workers and Test rounds are the parameters set up in the experiment. The number of transactions has been kept increasing starting from 1000 to 50000. Throughout the experiment the rate control is set to be fixed, the number of workers are kept at 50 and the test is carried for update and query operations.

**4.2. Impact on Success Rate.** The experiment observed the impact on success rate for update and query operations and is depicted in Fig. 4.1 and Fig. 4.2. It is observed that the success rate for update operation is 100 percent for 1000 and 2000 transactions per second and it starts to decline as the number of transactions increased after 3000. It is also evident that the success rate of query operations does not get affected with the number of transactions.

**4.3. Impact on Throughput.** The impact on throughput on varying the number of transactions for update and query operations are calculated and are shown in Fig. 4.3 and Fig. 4.4. It is found that the throughput of the proposed system for 1000 transactions is 100 percent and the system is able to give above 90 percent for upto 30000 transactions in case of update operations. For query operations, it is found that the throughput is 100 percent for 1000, 2000, 3000 transactions. It is found to be above 90 percent for 5000, 10000 and 30000 transactions. It founds to decrease only after 40000 transactions.

Transaction	Success Rate of Query (%)
1000	100
2000	100
3000	100
5000	100
10000	99
20000	97
30000	97
40000	95
50000	91

Table 4.2: Impact on Success Rate

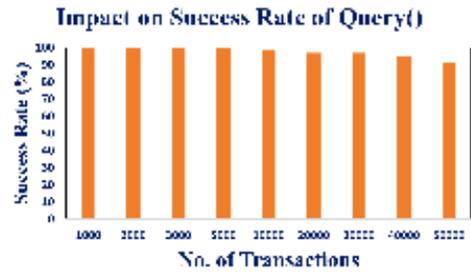


Figure 4.2: Impact on Success Rate of query()

Transaction	Throughput of Update (%)
1000	100
2000	100
3000	98
5000	98
10000	94
20000	92
30000	90
40000	80
50000	65

Table 4.3: Impact on Throughput

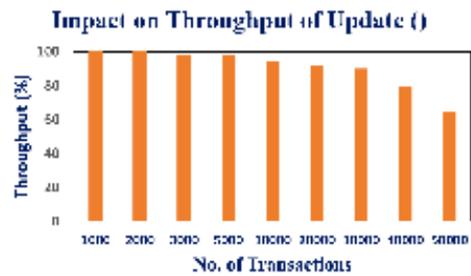


Figure 4.3: Impact on Throughput of update ()

Transaction	Throughput of Query (%)
1000	100
2000	100
3000	100
5000	97
10000	96
20000	93
30000	91
40000	83
50000	80

Table 4.4: Impact on Throughput

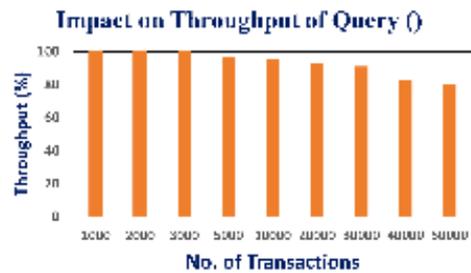


Figure 4.4: Impact on Throughput of query ()

Transaction	Latency of update()
1000	8
2000	16
3000	19
5000	21
10000	29
20000	34
30000	42
40000	70
50000	92

Table 4.5: Impact on Latency

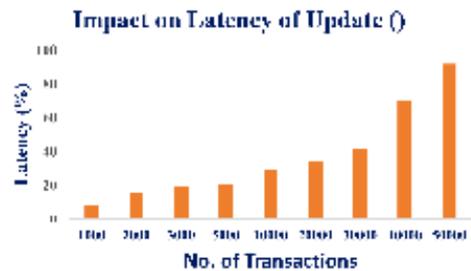


Figure 4.5: Impact on Latency of update ()

Transaction	Latency of query()
1000	3
2000	5
3000	12
5000	17
10000	21
20000	23
30000	28
40000	38
50000	43

Table 4.6: Impact on Latency

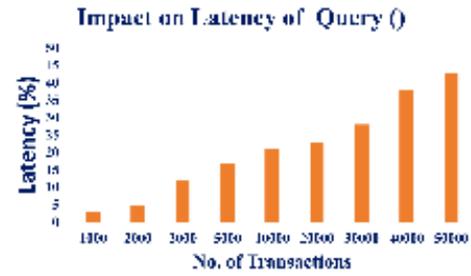


Figure 4.6: Impact on Latency of update ()

**4.4. Impact on Latency.** The latency for update and query transactions are observed and is found to be increasing in case of update transactions after 30000 transactions. It is also evident that the transaction latency is found to be high after 40000 transactions in case of query transactions. Fig. 4.5 and Fig. 4.6 illustrates the impact of number of transactions on latency in update () and query () operations.

**5. Conclusion and Future Work.** This work elaborates the significance of blockchain technology in various application areas. It envisions that blockchain stands as the one-stop solution for security and privacy issues. It also highlights that the blockchain technology in spite of its numerous advantages, suffers from scalability issue. This work proposed a channel based clustered sharding approach as the solution for scalability problem in blockchain. The system has been tested using the Hyperledger caliper, a blockchain benchmark tool for performance analysis of Hyperledger fabric framework. It is found that the proposed channel based clustered sharding system perform ten times faster than the system without sharding. It is also observed that the success rate and throughput begins to decrease and latency begins to increase once the system crossed 30000 transactions. It can be concluded that the channel based clustered sharding system performs ten times faster than the system without sharding. In the future, sharding can be used to store ledger data in the blockchain.

## REFERENCES

- [1] JANI, S, *The Emergence of Blockchain Technology & its Adoption in India.*,(2019)
- [2] NAKAMOTO, S, *Bitcoin: A peer-to-peer electronic cash system. Decentralized business review.* (2008).
- [3] DANNEN, C, *Introducing Ethereum and solidity* , Berkeley: Apress, (Vol. 1, pp. 159-160). (2017).
- [4] SABRY, S. S., KAITTAN, N. M., & MAJEED, I. , *The road to the blockchain technology: Concept and types.* , . Periodicals of Engineering and Natural Sciences., 7(4), 1821-1832. (2019).
- [5] AKCORA, C. G., GEL, Y. R., & KANTARCIOGLU, M. , *LU-Blockchain networks: Data structures of Bitcoin, Monero, Zcash, Ethereum, Ripple, and Iota*, Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 12(1), e1436.(2022).
- [6] RANKHAMBE, B. P., & KHANUJA, H. K. , *A comparative analysis of blockchain platforms–Bitcoin and Ethereum*, 5th international conference on computing, communication, control and automation (ICCUBEA) (pp. 1-7). IEEE., (2019)
- [7] PADMAVATHI, U., & RAJAGOPALAN, N, *Concept of blockchain technology and its emergence.* , In Research Anthology on Convergence of Blockchain, Internet of Things, and Security (pp. 21-36). IGI global.
- [8] YANG, D., LONG, C., XU, H., & PENG, S, *A review on scalability of blockchain.* , Proceedings of the 2020 the 2nd International Conference on Blockchain Technology (pp. 1-6).(2020, March).
- [9] KHAN, D., JUNG, L. T., & HASHMANI, M. A. , *Systematic literature review of challenges in blockchain scalability*, Applied Sciences, 11(20), 9372. (2021)
- [10] NASIR, Q., QASSE, I. A., ABU TALIB, M., & NASSIF, A. B. , *Performance analysis of hyperledger fabric platforms*, Security and Communication Networks, 2018.
- [11] AMIRI, M. J., AGRAWAL, D., & EL ABBADI, A. , *On sharding permissioned blockchains.* , IEEE International Conference on Blockchain (Blockchain) (pp. 282-285). IEEE. (2019).
- [12] MUTHUKUMARAN, V., V. VINOTH KUMAR, ROSE BINDU JOSEPH, MERAM MUNIRATHNAM, I. S. BESCHI, AND V. R. NIVEDITHA, *Efficient Authenticated Key Agreement Protocol for Cloud-Based Internet of Things*, In International Conference on Innovative Computing and Communications: Proceedings of ICICC 2022, Volume 3, pp. 365-373. Singapore: Springer Nature Singapore, 2022.
- [13] KOTI, MANJULA SANJAY, V. MUTHUKUMARAN, A. DEVI, V. RAJALAKSHMI, S. SATHEESH KUMAR, AND V. VINOTH KUMAR, *Efficient Deep Learning Techniques for Security and Privacy in Industry*, Cyber Security and Operations Management for Industry 4.0, October 27, 2022, 2–32.

- [14] KUMAR, VINOTh, V. R. NIVEDITHA, V. MUTHUKUMARAN, S. SATHEESH KUMAR, SAMYUKTA D. KUMTA, AND R. MURUGESAN, *A quantum technology-based lfi security using quantum key distribution*, In Handbook of Research on Innovations and Applications of AI, IoT, and Cognitive Technologies, pp. 104-116. IGI Global, 2021.
- [15] LUU, L., NARAYANAN, V., ZHENG, C., BAWEJA, K., GILBERT, S., & SAXENA, P, *A secure sharding protocol for open blockchains.* , Proceedings of the 2016 ACM SIGSAC conference on computer and communications security (pp. 17-30).
- [16] ZAMANI, M., MOVAHEDI, M., & RAYKOVA, M, *Rapidchain: Scaling blockchain via full sharding.* , In Proceedings of the 2018 ACM SIGSAC conference on computer and communications security (pp. 931-948).
- [17] AIYAR, K., HALGAMUGE, M. N., & MOHAMMAD, A. , *Probability distribution model to analyze the trade-off between scalability and security of shardingbased blockchain networks.* , IEEE 18th Annual Consumer Communications & Networking Conference (CCNC) (pp. 1-6). IEEE. (2021).
- [18] KEDZIORA, M., PIEPRZKA, D., JOZWIAK, I., LIU, Y., & SONG, H. , *Analysis of segregated witness implementation for increasing efficiency and security of the Bitcoin cryptocurrency*Journal of Information and Telecommunication, 7(1), 44-55.
- [19] C.N.S.VINOTH KUMAR, VASIM BABU M, NARESH R, LAKSHMI NARAYANAN K, BHARATHI V , , *Real Time Door Security System With Three Point Authentication*” , 4th International Conference on Recent Trends in Computer Science and Technology (ICRTCSST), IEEE Explore , May 2022. DOI: 10.1109/ICRTCSST54752.2022.9782004.
- [20] XU, Z., HAN, S., & CHEN, L, CUB, a consensus unit-based storage scheme for blockchain system. IEEE 34th International Conference on Data Engineering (ICDE) (pp. 173-184). IEEE.(2018).
- [21] DAI, X., XIAO, J., YANG, W., WANG, C., & JIN, H , *Jidar: A jigsaw-like data reduction approach without trust assumptions for bitcoin system*, IEEE 39th International Conference on Distributed Computing Systems (ICDCS) (pp. 1317-1326). IEEE. (2019).
- [22] LEWENBERG, Y., SOMPOLINSKY, Y., & ZOHAR, A. , Inclusive block chain protocols. In Financial Cryptography and Data Security, 19th International Conference, FC 2015, San Juan, Puerto Rico, January 26-30, 2015, Revised Selected Papers 19 (pp. 528-547). Springer Berlin Heidelberg.(2015).
- [23] SOMPOLINSKY, Y., & ZOHAR, A, *Phantom. IACR Cryptology* , ePrint Archive, Report 2018/104.(2018).
- [24] POON, J., & DRYJA, T. , *The bitcoin lightning network: Scalable off-chain instant payments.*, (2015).
- [25] DECKER, C., & WATTENHOFER, R , *A fast and scalable payment network with bitcoin duplex micropayment channels.* , In Stabilization, Safety, and Security of Distributed Systems: 17th International Symposium, SSS 2015, Edmonton, AB, Canada, August 18-21, 2015, Proceedings 17 (pp. 3-18). Springer International Publishing.
- [26] MILLER, A., BENTOV, I., BAKSHI, S., KUMARESAN, R., & MCCORRY, P. , *Sprites and state channels: Payment networks that go faster than lightning*, Financial Cryptography and Data Security: 23rd International Conference, FC 2019, Frigate Bay, St. Kitts and Nevis, February 18–22, 2019, Revised Selected Papers (pp. 508-526). Cham: Springer International Publishing. (2019).
- [27] LU, N., ZHANG, Y., SHI, W., KUMARI, S., & CHOO, K. K. R. , *A secure and scalable data integrity auditing scheme based on hyperledger fabric*, Computers & Security, 92, 101741.(2020).
- [28] HONAR PAJOOH, H., RASHID, M. A., ALAM, F., & DEMIDENKO, S. , *Experimental Performance Analysis of a Scalable Distributed Hyperledger Fabric for a Large-Scale IoT Testbed*, Sensors, 22(13), 4868. (2022)

*Edited by:* Polinpapilinho Katina

*Special issue on:* Scalable Dew Computing for Future Generation IoT systems

*Received:* Jul 24, 2023

*Accepted:* Dec 20, 2023



## FEATURE EXTRACTION AND CLASSIFICATION OF GRAY-SCALE IMAGES OF BRAIN TUMOR USING DEEP LEARNING

KONDRA PRANITHA\* AND NARESH VURUKONDA †

**Abstract.** Deep Learning using CNN plays a paramount role for the classification methods applied on medical image data. With a crucial role in accurate diagnosis, treatment planning and patient management for medical and healthcare systems, CNNs won accolades in the Deep Learning research. As simple the learning model so precise are the results for decision making. The proposed Sequential model of CNN is built with Parametric ReLU with the values aligned to geometric mean, attains a specific goal of tumor classification. The additional support of ground-truth aid in deciding the shape and severity of tumor in the Grayscale MRI of brain tumor. The simple Sequential model, although a minimal version has proved achieved significant classification goals using the GMP-ReLU. Comparative results with variants of ReLU have been charted in this article standing with the proof of consistent classification model with parametric-ReLU. The proposed design is conducted on images from Kaggle and a model is trained (classifier is built), which can be considered as ideal filter for all the benchmark images. The accuracy of proposed design is considerably improved compared to normal ReLU up to 89.214%

**Key words:** Deep learning, ground-truth based classification, parametric ReLU

**1. Introduction.** Brain is the most critical and vital organ in the human body. The root cause of dysfunction of the brain is the brain tumor. Excess of cells growing in the inter- and intra-circulatory regions of the brain causes lumps and muscular lesions called tumors. Tumors grow by consuming the nutrient inputs considered taking for the sake of body health. The tumors are located manually by expert clinicians and doctors in the human body using MRI images of the subjects, often causing inaccuracies and cumbersome in tracing, which takes more time. Brain cancer is caused by tumors which is critical and even causes death [1-3]. Therefore it shall be considered screened early in order to initiate remedial medicinal care and prevent well before growth of tumor is advanced. Classification helps in identifying the stages of the growth in tumors that cause deadly brain cancer [4-7]. Clinical diagnosis has more challenging issues related to tumor and cancer classification and identification. Medical Imaging introduces the medical image study methodology, where the experts can identify the crust of the tumors, edema or unorganized growths stemming in the inter-circulatory regions of the human body [8-10]. There are many imaging technologies invented since ages like computed tomography, positron emission tomography, magnetic resonance imaging and ultrasound. Even more complex imaging techniques were also developed to interact with the various sites of the human body[11]. Many medical institutions of various levels are applying the medical imaging technology which helps experts in visually diagnosing the internal regions anatomical and pathological strictures in the human body. Medical imaging protocols are developed envisioning the various types of diagnostic needs with different modalities of images.

The key stone in the gamut of mitigating with the patients affected with brain tumor is only early detection. Early detection reduces life and danger helps increased hope of survival possibilities. More than 90% of subjects recovered through early detection, whereas factor for not choosing is cost and availability of resources [12-14]. The best economic method would be detection and automation of image analysis through computer aided detection and diagnosis systems[15](CADDs).

The most popular method of imaging in the contemporary procedures is magnetic resonance imaging, which is more specifically used in diagnosis of brain cancer. As the brain cancer is arbitrary in size, location and affects type of the tumor, it is unpredictable for the conventional practitioners to guard on the eruption of cancer

---

\*Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, AP 522502, India ([kondra.pranitha@gmail.com](mailto:kondra.pranitha@gmail.com))

†Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, AP 522502, India ([naresh.vurukonda@kluniversity.in](mailto:naresh.vurukonda@kluniversity.in))

in the brain [16]. General metabolic flows in the brain are blocked by the tumors occurred by compressing the cells in the periphery of certain points which raises pressure, causing symptoms like nausea, imbalance walking and headache [17]. Such kind of inter-circulatory operations inside the brain cannot be traced with ultrasound or X-ray, where the malignity is location specific. Based on the examination required for the diagnosis, CT or MRI scanning is preferred to trace the foreign masses accurately. Most popularly used strategy for brain tumors is MRI which images the masses and the surroundings. Although, conventional procedures of MRI and CT support medical experts to review, as the MR images are high-definition images, rigid, non-ionizing and lasts considerably for longer time.

**2. Literature Survey.** Machine Learning has several steps to go through before an inference or knowledge is drawn from the data sets, such as: data augmentation, pre-processing, sampling, feature-detection, extraction, reduction and classification [18]. Feature extraction is the very important activity of the machine learning algorithms, in problems like classification, where the accuracy of the classification lies in extracting more positive features. Feature extraction considered as a low-level process, uses first-order statistics (like mean, median, standard deviation and skewness) are used in the low-level feature extraction [19-22]. Second order statistics such as shape, Gabor features and gray-level co-occurrence matrix were used to understand the formation of the foreign masses.

First order and second order statistics are employed for linear least square variant of support vector machine, which is used to develop the binary classifier, to classify malign or benign categories of brain MRI images [23]. Low-level features describe the tumor images efficiently, in spite of their representational efficiency due to similarity in appearances and anonymity of texture, shape and size, therefore gray level co-occurrence matrix (GLCM) [24-25] is applied for tumor identification and classification.

Hence, the traditional machine learning methods have problem in feature extraction stage, feature extraction need to be handcrafted based on the information available and methods applied, and there are more chances of human errors [26-27]. At this juncture, it is very essential to develop automated methods to refine the method of extraction of features using various types and levels of statistical information. An MRI image of brain has many mentions of details which are not in the texts and consensus of brain tumor literatures. In histopathology, experts can progressively identify and analyze certain impressions of features in MRI images [28-30]. Transformational reactions such as apoptosis and mitosis are not precisely studied in the conventional methods by the pathological experts, which prone to negligible errors, which even may cause morphologically a severe health menace. Histopathologists read certain features which has diagnostic leads, which are confined to limited subjectivity [31]. Other features which are found as peculiar and new to pathological analyses shall be extracted by data-driven approaches. A data-driven approach can grade the symptoms with considerable efforts of automated learning mechanism. ANN, SVM and Random Forests are allied methods that form into a conglomerate of analytical learning on the images, therefore the image analysis would pave into data driven learning and quantitative learning with salient pathological features.

Detecting the varying shape of the alien growth in the internal regions of the brain is not only relied on pattern recognition, where other non-structural properties also are influenced. For such situations, the parts that grow shall be examined as whether healthy or not. A brain atlas shall be studied for the various types of tissues in the internal regions of the brain, to assess the lesions are healthy or not. Symmetry of brain is also another important study where abnormal regions are detected that violates the symmetry.

During the study and classification of brain tumor, the reflections of the signals determine the type of tissue whether a tumor or not. The same signal intensities as that appear for a normal tissue appears for tumors which are categorized as isointense [27]. The isointense to hypo-intense properties of the tissues in the human brain associated with perilesional edema confuses the classification of brain tumor. Primary and secondary tumors such as meningioma (MEN), and metastasis (MET) respectively are detected using T1-weighted images of MRI for classification.

**3. Brain Tumor Gray-scale MRI.** One among the non-invasive techniques widely applied for diagnosis and monitoring of brain activities and tumors is MRI. Gray-scale MRI images provide valuable information about the tumor and surrounding tissue [13]. The resolution of these images can vary based on the acquisition protocol, imaging equipment, and specific clinical requirements. Some common resolutions for brain tumor gray-scale MRI images are:

*Low resolution:* Lower resolution images typically have a larger voxel size, which means each voxel represents a larger area in the scanned tissue. This can result in images with a resolution of around 256x256 or 128x128 pixels. While low-resolution images can be acquired quickly and require less computational resources for processing, they may lack the fine details needed for accurate diagnosis and treatment planning.

*Medium resolution:* Medium resolution images offer a balance between speed of acquisition and detail, with voxel sizes that are smaller than low-resolution images but larger than high-resolution images. These images often have resolutions around 512x512 pixels. Medium resolution images are commonly used in routine clinical practice, as they provide sufficient detail for most diagnostic purposes.

*High resolution:* High-resolution MRI images have small voxel sizes and capture finer details of the brain tissue, which can be crucial for certain applications, such as surgical planning or research. These images may have resolutions of 1024x1024 pixels or higher. High-resolution imaging can be time-consuming and require more advanced imaging equipment and computational resources for processing and analysis.

*Super-resolution:* Super-resolution is an advanced image processing technique that can be used to enhance the resolution of MRI images beyond the limits of the imaging equipment. Multiple image datasets with varied resolutions are combined from low-resolution images and super-resolution algorithms generate images with reasonable resolutions better comparable with high-resolution MRI. This can be particularly useful for applications where high-resolution images are needed but are difficult to acquire directly due to time or hardware constraints.

The choice of resolution for brain tumor gray-scale MRI images depends on the specific clinical scenario, diagnostic requirements, and available imaging equipment. In general, higher resolution images provide more detail and can potentially lead to more accurate diagnoses and treatment plans, but they also require more time to acquire and process.

**4. Proposed Work.** Brain tumor image analysis and segmentation is a subsidiary field of medical image analysis. From the research consensus in the recent decade, computer vision algorithms have profound contribution supporting diagnosis, treatment and monitoring the behavioral aspects of brain when affected by tumors. Much of the key features related to brain tumor such as size, location, type, grade, shape and boundaries, enhancing patterns, molecular features shall be understood during analyses shown in Fig.4.1.

**4.1. Method of feature selection.** Feature Selection has predominant role in classification and other prediction techniques. Feature selection is a preprocessing activity in deep learning for classification. Swarm based algorithms consists of two important characteristics which commonly prevail to demonstrate the meta-heuristic mechanism. Intensification and diversification are the two characteristics in swarm algorithms which do exploitation and exploration in the search space. Intensification aims at finding the best solution with respect to the current generation population in swarm optimization, whereas diversification is exploring the properties in the search space in stochastic orders, useful to evaluate the effectiveness of the algorithm. Most machine learning and deep learning algorithms employ meta-heuristic algorithms for feature selection. Meta-heuristic algorithms of swarm type are most suitable for the detection and identification of features in medical images. The phenomena of biological foraging behavior of animals, insects and sea-beings are most suitably applied for swarming in the feature selection of Fig.4.2.

**4.2. Ground Truth images and their importance.** The process of segmentation of brain tumor images from MRI is dynamic with respect to feature detection [46][51]. The process is vital in detecting features of tumors; delve in planning of treatment and to extract some findings different from healthy subjects. This is a challenging process, where varieties of methods are proposed in the research consensus. Manually rating the process of finding the standard validation in detecting the tumor features is always compared with the systematic or automated methods. The manual process of segmentation hinders with reliability, interpretation and reproducibility, more particularly at different regions employ different methods. Therefore, a ground truth needs to be prepared. a true ground truth is needed in segmentation of images manually. The importance of ground truth in detection and extraction of the known tumor features is more helpful. From the manual processes of segmentation of edema, several samples and estimates are considered to prepare the true ground truth.

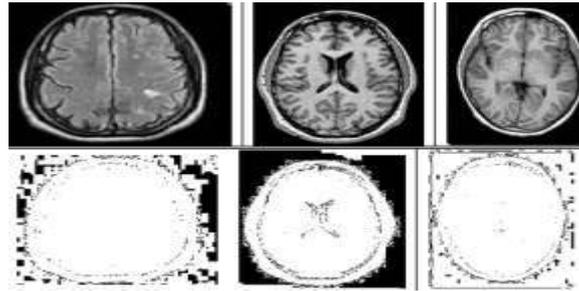


Fig. 4.1: Kaggle Images of Brain Tumors with ‘no tumor’ and the respective ground truths

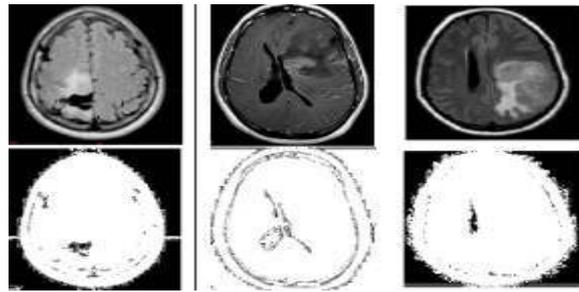


Fig. 4.2: Kaggle Images of Brain Tumors with ‘yes tumor’ and the respective ground truths

As the anatomy of brain tumor is complex, locating the points and positions of tumor requires intervention of experience of specialists. Characteristics of brain tumor and the detailed knowledge of brain functions and anatomy are essential for medical professionals including neurologists, oncologists and neuro-radiologists.

The following are the ground truth images of brain tumor data sets from Kaggle.

Ground truth images are annotated and labeled from the collective sources, where the reference information about the images is composed. Applied popularly in the deep learning techniques associated with classification, segmentation and object detection. Using ground truth, the performance of the algorithms can be evaluated with greater accuracy for the target sites in the medical images of brain tumor.

Ground truths for brain tumor images can be obtained by various methods such as manual segmentation, semi-automated segmentation, consensus labeling, synthetic or simulated data, which helps in developing and evaluating the performance of reliable computer vision algorithms.

**4.3. Composition of CNN for selecting features using ground truths.** The well-known fact about CNNs is it is a powerful class of deep learning model, which proves to be with exceptional performance particularly in computer vision research, such as segmentation, detection and recognition. When using CNNs for selecting features based on ground truths, the network is designed to learn meaningful representations from the input data, which can then be used for the given task. Here is a general outline for designing a CNN to select features using ground truths:

**Input Layer:** The raw image data sets are usually preprocessed and augmented considering the key parameters such as height, width and color channels and conversion into arrays.

**Convolutional Layers:** The building blocks of CNN, consists of various types of filters, kernels and striding mechanism performed through a convolution operation. Filters and kernels learn local features such as textures, corners and edges. **Activation functions:** At each convolution layer an activation takes place with a specific chosen function, to introduce the non-linearity, such as ReLU to substitute the equivalent values or chosen values for improvising the learnability.

**Pooling Layers:** Spatial dimensions of the images are converted to feature maps in pooling layers. The

computational complexity is reduced by overcoming the problems of overfitting. Max-pooling is the ideal process which considers the maximum values from a region in the feature map.

**Fully connected layers:** Feature maps in the form of arrays are flattened into a single vector after the pooling layer. Global features and relationships among different images are enabled to learn.

**Output layer:** The output layer is connected to the last fully connected layer and is responsible for producing the final feature vector or classification scores. Depending on the task, different activation functions are used, such as the softmax function for multi-class classification or sigmoid for binary classification.

**Loss Function:** The difference between the predictions and the ground truth labels are drawn, considering the cross-entropy loss during classification.

**Optimization:** An optimizer is used to revise the update the weights as per the learning requirements during training while minimizing the loss function.

**Regularization techniques:** To prevent overfitting and improve generalization, regularization techniques such as dropout, weight decay, or data augmentation can be applied. When designing a CNN for selecting features using ground truths, the network architecture, hyperparameters, and training strategy should be tailored to the specific task and dataset. This may require experimenting with different architectures, layer configurations, and learning rates to find the best combination that produces the most accurate and robust model.

**4.4. Various Activation Functions used Feature Selection using Ground Truth.** Activation functions play an essential role in Convolutional Neural Networks (CNNs) by introducing non-linearities that enable the network to learn complex patterns from the input data. There are several types of activation functions used in CNNs for feature selection based on ground truths. Some of the most common activation functions are: Rectified Linear Unit (ReLU): ReLU is the most widely used activation function in CNNs. It introduces non-linearity by setting all negative values in the input to zero, while keeping positive values unchanged. The simplicity of ReLU helps with faster training and reduced likelihood of vanishing gradients shown in Equ.4.1.

$$f(x) = \max(0, x) \quad (4.1)$$

**Leaky Rectified Linear Unit (Leaky ReLU):** Leaky ReLU is a modified version of the ReLU function that allows for a small, non-zero gradient when the input is negative. This helps to mitigate the "dying ReLU" problem, where some neurons become inactive and stop learning during training shown in Equ.4.2.

$$f(x) = \max(\alpha x, x) \quad (4.2)$$

where  $\alpha$  is a small constant (e.g., 0.01)

**Parametric Rectified Linear Unit (PReLU):** PReLU is another variant of ReLU that generalizes Leaky ReLU by learning the slope of the function for negative input values during training. Equ.4.3 enables the model to adapt better to the specific characteristics of the dataset.

$$f(x) = \max(\alpha x, x), \quad (4.3)$$

where  $\alpha$  is a learnable parameter

**Exponential Linear Unit (ELU):** ELU is an activation function that smooths the transition between positive and negative values by using an exponential curve. This smoothness helps to mitigate the vanishing gradient problem and speeds up training shown in Equ.4.4.

$$f(x) = x \text{ if } x \geq 0, \text{ else } \alpha(\exp(x) - 1) \quad (4.4)$$

**Scaled Exponential Linear Unit (SELU):** SELU is an activation function designed specifically for feed-forward neural networks with normalized weights and activations. When used with proper weight initialization and normalization, SELU can help the network to self-normalize, improving training stability and performance.

$$f(x) = \lambda x \text{ if } x > 0, \text{ else } \lambda \alpha(\exp(x) - 1) \quad (4.5)$$

In Equ.4.5,  $\lambda$  and  $\alpha$  are predefined scaling factors

Table 5.1: Contrast Ratio

Contrast Ratio	RGB Values			Number of Images	Sample Size
	R	G	B		
	255	128 – 178	0 – 102	Total : 155	Avg. Size: 15
16	255	128	0	30	12
18	255	136	14	40	8
20	255	144	28	50	10
21	255	152	42	70	12
22	255	160	58	85	15
24	255	166	72	115	8
26	255	172	86	130	8
28	255	178	102	155	12

**Hyperbolic Tangent (tanh):** The tanh function is a scaled and shifted version of the sigmoid function that maps the input values to the range between -1 and 1. This activation function is less commonly used in CNNs due to the vanishing gradient problem, but can still be effective in certain cases shown in Equ.4.6.

$$f(x) = (exp(2x) - 1)/(exp(2x) + 1) \quad (4.6)$$

**Sigmoid:** The sigmoid function maps input values to the range between 0 and 1. While it is not commonly used as an activation function in hidden layers of CNNs due to the vanishing gradient problem, it is still useful in output layers for binary classification tasks.

$$f(x) = 1/(1 + exp(-x)) \quad (4.7)$$

Equ.4.7 gives the right activation function for a CNN depends on the specific task, dataset, and network architecture. In practice, ReLU and its variants (Leaky ReLU, PReLU) are the most popular choices due to their simplicity and effectiveness in most scenarios.

**5. Results and Discussion.** Synthetic samples of gray-scale brain tumor images are collected as input for the experimentation. The Kaggle data set is proposed with 253 images where 155 containing the symptoms of brain tumor and 98 images without any symptoms. The consideration of generating synthetic image data sets is to make the sample sizes bigger in order to provide larger scope for deep learning. Around 800 synthetic images are generated using Dezgo [www.dezgo.com], where the consistent availability of symptoms is unpredictable and suitable as an unknown input for the experiment. Most of the braintumor images drawn from the dataset contain pixels considering the hue values of 25 to 50, and saturation and luminosity values set up at constant 76% and 62% respectively in the regions notified as tumors. The regions look as blurry non-geometric areas which contain no crispy borders to locate the size and shape. Whether images contain features of tumor or not, it is very effortful to identify based on the colors, if color images are provided, therefore a Contrast Limited Adaptive Histogram Equalization method for the selected RGB density values of such areas are considered to supply as input the classification process to ascertain the presence of proliferative components. From the selected images containing suitable densities of pixels, as shown in the following table, a Sequential Model of CNN with kernel values as densities are iterated using various types of activation functions. The table shows number of images that arrive satisfactorily in the experiment for each activation function. Given with various levels of selected contrast ratios (drawn from the observations of the experiment), the range of R, G, B values are identified using RGB histograms (dcode.fr), for the images with tumor given in the Kaggle data sets. Number of images for each average sample size of 15 from the total number of images with tumor are noted in Table.5.1.

Typically an image histogram is a statistical graph, which represents the colors in the x-axis and distribution of RGB colors and Luminosity of image on the y-axis. A gross statistic of the image histogram represents a gray-scale version of luminosity and it is calculated for each pixel with the standard formula  $0.2126 * R + 0.7152 * G + 0.0722 * B$ .

Table 5.2: Classified Samples containing tumors with respect to implemented activation function

Description of Sample Selection (Contrast Ratio,Color after contrast)	No. of Images with tumor	Sample Size	Images in Classified Samples containing tumors with respect to implemented activation function				
			Sigmoid	Tanh	ReLU	Leaky ReLU	GMP-PreLU
Contrast Ratio : 16 Color : RGB(255, 128, 0)	30	12	8	9	8	7	10
Contrast Ratio : 21 Color : RGB(255, 152, 42)	60	12	10	10	9	8	11
Contrast Ratio : 22 Color : RGB(255, 160, 58)	65	15	8	10	8	7	12
Contrast Ratio : 26 Color : RGB(255, 172, 86)	70	8	5	8	6	5	7
Contrast Ratio : 28 Color : RGB(255, 178, 102)	75	12	8	10	8	7	12

The cloudy areas in the image represent tumor, the said color densities with contrast of 1% to 100% are considered in as a geometric progression shown in Table.5.2. The reason of using geometric progression in the experiment is the nature of the exponential spread style of the colored pixels. A pixel (p) is selected as a seed point of the random walk in the brain tumor image; the next pixel is detected with the same color with a distance(d). Thus, the infinite sum of the pixels is computed in geometric progression as in Equ.5.1.

$$\sum_{k=0}^{\infty} (p.d^k) = p \frac{1}{1-d} \tag{5.1}$$

For a sample window of the pixel values a series of contrast ratios accounted for the computation of the Geometric Progression sum, which may be considered as a parameter for the PReLU as first case in Table.5.3. Alternative case for Parametric ReLU is computed with value of  $\alpha$  determined from a predefined value of pixel intensity with background knowledge, which is equivalent to the standard deviation of selected contrast ratios i.e., 16, 21, 22, 26, 28 is the value called ‘p’ which is replaced with  $\alpha$ . The geometric progression of the min and max values of the selected contrast ratios of samples in an instance of observation is considered as an ideal value for the parameter shown in Table.5.4. Mean = 22.6, Sum of Contrast Ratios = 113 and Variance = 17.44

$$StandardDeviation(16, 21, 22, 26, 28) = \sqrt{1/N \cdot \sum_{i=1}^N (x_i - \mu)^2} = 4.1761226035642 \tag{5.2}$$

Margin of Error (Confidence Interval):

The distribution of the sampling mean is mostly as normal distribution. In this experiment, the standard error of the mean is computed as in Equ.5.3.

$$\sigma_x = \frac{\sigma}{\sqrt{numberofcontrastratio}} = 1.867618804 \tag{5.3}$$

Thus from the above observation of an instance of experimentation, the value of parameter ‘p’ is still corrected to  $p = sd+sem$ , where, if the confidence is considered as 95% of samples possess tumors  $p1 = sd+sem$ , for less than 95% samples  $p2 = sdsem$ .

Table 5.3: Images in Classified Samples containing tumor with respect to implemented activation function GMP-PReLU

Description of Sample Selection (Contrast Ratio, Color after contrast)	No. of Images with tumor	Sample Size	Images in Classified Samples containing tumor with respect to implemented activation function GMP-PReLU		
			p1	p2	average
Contrast Ratio : 16  Color : RGB(255, 128, 0)	30	12	12	8	10
Contrast Ratio : 21  Color : RGB(255, 152, 42)	60	12	12	10	11
Contrast Ratio : 22  Color : RGB(255, 160, 58)	65	15	14	10	12
Contrast Ratio : 26  Color : RGB(255, 172, 86)	70	8	8	6	7
Contrast Ratio : 28  Color : RGB(255, 178, 102)	75	12	12	12	12

Table 5.4: Convolution Filters used in the Sequential Model of CNN

Layer(type)	Output shape	Param#
Conv2d(conv2D)	(None,30,30,32)	896
Max.Pooling2d(Max.pooling 2D)	(None,15,15,32)	0
Conv2d(Conv2D)	(None,13,13,64)	18496
Max.Pooling2d(Max.Pooling2D)	(None,6,6,64)	0
Conv2d(conv2D)	(None,4,4,64)	36928
Flatten(Faltten)	(None,1024)	0
Dense(dense)	(None,64)	656000
Dense(dense)	(None,10)	650

Detecting features is challenging as they are not found in specific spatial positions in the image. The convolution is performed on the various spatial areas of the image to determine whether it is a tumor Table.5.5.. In contrast to traditional neural network, the image has to transformed into a tensor rather into a vector in order to proceed with convolutions. A convolution filter is prepared, as a thumb rule with a size of 3x3 or 5x5. The following figure shows the 3x3 convolution filters:

The objective of the present application of Convolutional Neural Network is classification of the images with tumor from Kaggle repository in Table.5.6. The Keras Sequential API is primarily used and further up kept with modifications in convolution filters, ReLU and levels of layers.

```

model = models.Sequential()
model.add(layers.Conv2D(32, (3, 3), activation = 'relu', input_shape = (32, 32, 3)))
model.add(layers.MaxPooling2D((2, 2)))
model.add(layers.Conv2D(64, (3, 3), activation = 'relu'))
model.add(layers.MaxPooling2D((2, 2)))

```

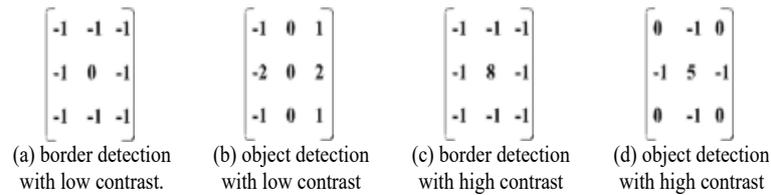


Fig. 5.1: Convolution Filters used in the Sequential Model of CNN

```
model.add(layers.Conv2D(64, (3, 3), activation = 'relu'))
```

The above lines of code are supported by a stack of 2D convolution layers and max pooling layers. The image input is prepared with parameters of width, height and the color channel irrespective of the batch size. A batch is the size / one chunk of input given to the Convolution Neural Network. The CNN processes the input with 32, 32, 3 as the width, height and the color channel respectively. The original size of the brain tumor images of the Kaggle datasets is 225 x 225, where a regular patch is extracted from the images of size 32x32x3. The first layer is given an input of a batch of size 16 images with 32x32x3 as input specifications. Tensors representing the width and height as dimensions of the images are the output from the convolution layer, where the images tend to shrink into deeper layers of the convolutional neural network model. The first argument controls the number of output of the channels (tentatively 32), as specifically width and height shrinks, more output channels are added to each convolution layer.

Tensor is the last output in the model, where it is fed into the convolution layer or into one or more dense layers, during classification, with the dimension of convolution as 4x4x64. As dense layers work on vectors, the tensors are converted into vectors, by means of flattening the tensor or unrolling the tensor, subsequently the dense layer is added at the top. The Kaggle datasets contains prospectively 2 to 4 classes; final dense layers of 2 or 4 in number are connected at the end of the convolutional network at 2 or 4 outputs. The following model summary demonstrate the CNN used in the experiment.

The network model proposed summarizes the model and the tensor (4, 4, 64) is flattened into vectors of shape (1024) with advanced application of two dense layers. Using adam optimizer of Keras and setting values of sparse categorical cross entropy in logit distribution the model is compiled and the accuracy and loss is traced in 10 epochs. The model is fit to the input and evaluated, for 10 epochs; the accuracy is obtained to 0.7192 and loss at 0.8475 with almost at average of 634 m/s per epoch. Further the model is evaluated for increasingly 15, 20, 50, 100, 150 epochs successively on the same inputs. The accuracy is obtained to 0.8792 and loss at 0.6432 with almost average of 714 m/s per epoch. In the proposed work, GMP-PReLU is implemented for better finding of images with dense clouds identified as tumors gray-scale images. The creation of GM-PReLU is a custom ReLU using Lambda layers in Tensorflow. An activation function can be created or edited using existing activation function. An example of implementation of custom ReLU is shown in the above code. The GMP-PReLU is introduced from standard samples of Kaggle data sets as cited previously. Selected samples of images containing severe features of tumors are considered and the pixel contrast ratio are computed. The pixel contrast ratios of the images are considered for the determination of the value of the 'p'.

As the parametric ReLU intervenes with an external input as a learnable parameter to regulate the learnability of ReLU, a value is computed consisting of characteristics of Geometric Mean, which improves the learning behavior in the classification model for Grayscale MRI of Brain tumor, exponentially compared with normal ReLU. Though, ReLU is the most popularly used activation function, which introduces non-linearity for all the negative values as zeroes, the linear identity of the positive values cannot draw suitable inferences of learning for all the images data with sparse information. Whereby, the activation function leads to null for most of the zero valued inputs in ReLU. To enhance such practical feasibility to overcome the flaws of learnability, Leaky ReLU can be worked out with the constant values of 0.01 or 0.001 for all the negatively assumed values in the sparse information of the image data sets. Whereas in the Parametric ReLU the constant values are instantaneously computed based on the population of gray-scale values for the image of the interest. Geometric

Table 5.5: Validation Parameters

Parameter	Computation
Accuracy	$\frac{TP+TN}{TP+TN+FP+FN}$
Error Rate	$\frac{FP+FN}{TP+TN+FP+FN}$
Positive Predict Value (PPV)	$\frac{TP}{TP+FP}$
Sensitivity	$\frac{TP}{TP+FN}$
Specificity	$\frac{TN}{TN+FP}$

Table 5.6: Classification with Sequential CNN using normal ReLU

No. of Samples	Observed Attempts		Cumulative Attempts		FPR	TPR	AUC
	Success	Fail	Success	Fail			
			0	0	1	1	0.074523
200	160	31	160	31	0.925477	0.972973	0.092901
250	205	36	365	67	0.829995	0.941587	0.105254
300	240	52	605	119	0.718211	0.896251	0.124398
350	298	57	903	176	0.579413	0.846556	0.139976
400	355	129	1258	305	0.414066	0.734089	0.11078
450	324	166	1582	471	0.263158	0.589364	0.071097
500	259	193	1841	664	0.142524	0.421099	0.034519
550	176	168	2017	832	0.06055	0.274629	0.010489
600	82	162	2099	994	0.022357	0.133391	0.002982
650	48	153	2147	1147	0	0	0
	2147	1147					0.76692

Mean of the trial values of positive values is calculated for the forthcoming values, if they are zeroes. The interpolative mechanism of employing values of Geometric Mean makes learning profitable. Deep Learning has ubiquitous methods for the problems of image classification, image annotation, image recognition and detection. The proposed model of CNN consists of 4 convolutional layers which are design with 16 feature maps. Thereby a Sequential Model of CNN could employ the ReLU with the said kind of parameters to avoid the saturation in learning. The feature maps after Max-Pooling is introduced with a kernel size of 2 x 2 and the normalization is directed by the normalization layers after each instance of Max-Pool layer, to enable faster convergence. Evaluating the performance of experiments is done by plotting the performance metrics of the observations in each instance of experiment. False Positive Rates and True Positive Rates are computed from all possible values of performance metrics and a curve is plotted. The curve between FPR and TPR denotes the gross variations of Sensitivity, Specificity and Accuracy achieved from the experiments. Wherefore, the experiments are conducted in two phases, that of a Sequential Model of CNN with normal characteristics of ReLU and GMPReLU, which is where the values obtained, are shown as ROC curve – Receiver Operating Characteristic curve. The following tables show tabulated data of the observations of the experiments of Sequential Model of CNN with ReLU and GMPReLU.

Evaluating the performance of experiments is done by plotting the performance metrics of the observations in each instance of experiment. False Positive Rates and True Positive Rates are computed from all possible values of performance metrics and a curve is plotted. The curve between FPR and TPR denotes the gross variations of Sensitivity, Specificity and Accuracy achieved from the experiments. Wherefore, the experiments are conducted in two phases, that of a Sequential Model of CNN with normal characteristics of ReLU and GMPReLU, which is where the values obtained, are shown as ROC curve – Receiver Operating Characteristic curve. The following tables show tabulated data of the observations of the experiments of Sequential Model of CNN with ReLU and GMPReLU.

The curve represents the performance of the experiments, where if the curve is closer to y-axis, the ex-

Table 5.7: Classification with Sequential CNN using GMP-ReLU

No. of Samples	Observed Attempts		Cumulative Attempts		FPR	TPR	AUC
	Success	Fail	Success	Fail			
			0	0	1	1	0.130485
200	226	4	226	4	0.869515	0.995859	0.152944
250	266	12	492	16	0.715935	0.983437	0.178291
300	314	14	806	30	0.534642	0.968944	0.186852
350	334	21	1140	51	0.341801	0.947205	0.211098
400	386	34	1526	85	0.118938	0.912008	0.036859
450	70	41	1596	126	0.078522	0.869565	0.032132
500	64	250	1660	376	0.04157	0.610766	0.014105
550	40	215	1700	591	0.018476	0.388199	0.005379
600	24	192	1724	783	0.004619	0.189441	0.000875
650	8	183	1732	966	0	0	0
	1732	966					0.94902

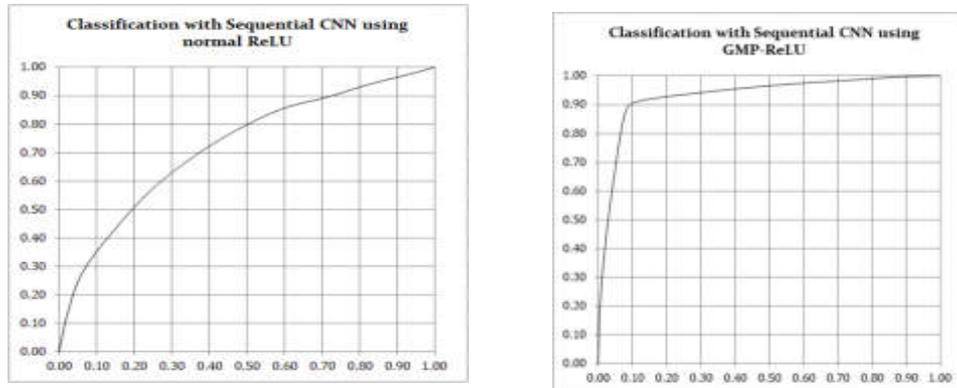


Fig. 5.2: The ROC showing the performance of the Sequential Model of CNN with ReLU and GMP-ReLU

periments exhibit a better fit. The AUC is also computed as the sum of values of product of the subsequent difference of FPR with the TPR is said to be achieved nearing to 1, indicating the maximum fit or better fit. Whereas, the value of AUC as 0.5 shows the indiscriminability of the model as success or failure of Table.5.7.

From observations of Fig.5.1, kernel shape and kernel size are progressively inverse. The reason is a very basic operational fundamental, as the kernel shape in first convolution passes, the resolution of the image is narrowed, in the sense the data is consolidated, thus by reducing size of kernel will increase proportionately. Number of convolutions is to narrow down the inferences obtained in the kernel, until a considerable inference is met; the convolutions are repeatedly performed. During the convolutions, the stride is considered to be 1 or 2, i.e., a unit matrix of one column and one row or a matrix with minimum countable columns and rows, which will be possibly 1 or 2. The stride indicates the number of kernel convolutions on the source image. Increase in the size of stride will reduce the resolution in the results as the number of convolutions becomes less in number.

As said above in the Sequential Model of CNN, application of filters means performing a convolution on the source. The size of the filter depends on the quality of comprehension required; size varies progressively in each iteration. The configuration of the framework consists of the pool size, stride size, computation on the pool, combining the results of computation from the convolutions by each filter. A filter is applied in convolutions on the source and completes all the computations of the source data, which is called as convolution layer. All the layers are combined into a tensor which is reshaped and flattened to a linear unit.

Ground truth based classification using convolution neural networks (CNNs) on gray scale images is a

common technique used in image recognition tasks. In this approach, a Sequential-CNN is trained using a data set of grayscale images that are labeled with the correct classes, known as the ground truth labels. The goal is to train the CNN to learn the features that are important for distinguishing between different classes of images. The process of training a CNN typically involves several steps, including data pre-processing, model design, and training. During the pre-processing step, the images are typically re sized and normalized to ensure that they are all the same size and have a consistent range of pixel values. This step is important for ensuring that the CNN can learn meaningful patterns from the images shown in Fig.5.2.

The model design step involves choosing the architecture of the CNN. Composition of multiple convolutional and pooling layers supporting convolution and pooling operations, which is typically followed by fully connected layers. The number and size of these layers will depend on the complexity of the problem being solved, and can be tuned through a process of experimentation and evaluation.

During the training phase, the labeled dataset is used in CNN for training, which is called backpropagation technique, where the weights are adjusted to minimize the variations among predicted labels and ground truth labels by the network. This process repeats as many epochs, or iterations, the network in the model achieve accuracy levels satisfactorily using training data.

A trained CNN for the specific input of training data sets can be conveniently used to classify new images of the training categories and at the final layer to determine the predicted class. Exclusive test data sets can be used to compare the performance of the network with predicted labels to the ground truth labels as a separate validation test.

**6. Conclusion.** In our experiment we have made an attempt to determine the tumor in Grayscale MRI of brain tumor. The training data is classified only for benign and malign. Meningioma as the primary and the metastases as secondary play rudimentary foal in the experiment. The synthetic data sets are used to build the training classifiers and then the training classifiers are applied on the real data sets in order to distinguish that with tumor or otherwise. Progressively, the iterations are performed in the experiment with varying sizes of kernel, in order to narrow the secondary tumors areas and improved resolutions of the inferences about primary tumors. The experiment is conducted on images from Kaggle and a model is trained (classifier is built), which can be considered as ideal filter for all the benchmark images. The accuracy of the model is considerably improved compared to normal ReLU up to 89.214%.

#### REFERENCES

- [1] Masood, Momina, Tahira Nazir, Marriam Nawaz, Awais Mehmood, Junaid Rashid, Hyuk-Yoon Kwon, Toqeer Mahmood, and Amir Hussain. "A novel deep learning method for recognition and classification of brain tumors from MRI images." *Diagnostics* 11, no. 5 (2021): 744.
- [2] Aggarwal, Ashwani Kumar. "Learning texture features from glcm for classification of brain tumor mri images using random forest classifier." *Transactions on Signal Processing* 18 (2022): 60-63.
- [3] Nawaz, Syed Ali, Dost Muhammad Khan, and Salman Qadri. "Brain tumor classification based on hybrid optimized multi-features analysis using magnetic resonance imaging dataset." *Applied Artificial Intelligence* 36, no. 1 (2022): 2031824.
- [4] Raza, Asaf, Huma Ayub, Javed Ali Khan, Ijaz Ahmad, Ahmed S. Salama, Yousef Ibrahim Daradkeh, Danish Javeed, Ateeq Ur Rehman, and Habib Hamam. "A hybrid deep learning-based approach for brain tumor classification." *Electronics* 11, no. 7 (2022): 1146.
- [5] Tummala, Sudhakar, Seifedine Kadry, Syed Ahmad Chan Bukhari, and Hafiz Tayyab Rauf. "Classification of Brain Tumor from Magnetic Resonance Imaging Using Vision Transformers Ensembling." *Current Oncology* 29, no. 10 (2022): 7498-7511.
- [6] Ahmed, Syed Thouheed, Vinoth Kumar, and JungYoon Kim. "AITel: eHealth Augmented Intelligence based Telemedicine Resource Recommendation Framework for IoT devices in Smart cities." *IEEE Internet of Things Journal* (2023).
- [7] Dhiman, Gaurav, V. Vinoth Kumar, Amandeep Kaur, and Ashutosh Sharma. "Don: deep learning and optimization-based framework for detection of novel coronavirus disease using x-ray images." *Interdisciplinary Sciences: Computational Life Sciences* 13 (2021): 260-272.
- [8] Babu, S. Sreedhar, and Polaiiah Bojja. "Abnormal Tumor Classification of MR Brain Images Using Deep Neural Networks." *Annals of the Romanian Society for Cell Biology* 25, no. 6 (2021): 4668-4680.
- [9] Guan, Yurong, Muhammad Aamir, Ziaur Rahman, Ammara Ali, Waheed Ahmed Abro, Zaheer Ahmed Dayo, Muhammad Shoaib Bhutta, and Zhihua Hu. "A framework for efficient brain tumor classification using MRI images." (2021).
- [10] Kang, Jaeyong, Zahid Ullah, and Jeonghwan Gwak. "MRI-based brain tumor classification using ensemble of deep features and machine learning classifiers." *Sensors* 21, no. 6 (2021): pp. 2222.
- [11] Irmak, Emrah. "Multi-classification of brain tumor MRI images using deep convolutional neural network with fully optimized

- framework." Iranian Journal of Science and Technology, Transactions of Electrical Engineering 45, no. 3 (2021): 1015-1036.
- [12] Khan, Hassan Ali, Wu Jue, Muhammad Mushtaq, and Muhammad Umer Mushtaq. "Brain tumor classification in MRI image using convolutional neural network." Mathematical Biosciences and Engineering (2021).
- [13] Badža, Milica M., and Marko Č. Barjaktarović. "Classification of brain tumors from MRI images using a convolutional neural network." Applied Sciences 10, no. 6 (2020): 1999.
- [14] Vinoth Kumar, V., V., Karthikeyan, T., Praveen Sundar, P. V., Magesh, G., & Balajee, J. M. "A quantum approach in life security using quantum key distribution." International Journal of Advanced Science and Technology 29 (2020): 2345-2354.
- [15] Gumaiei, Abdu, Mohammad Mehedi Hassan, Md Rafiul Hassan, Abdulhameed Alelaiwi, and Giancarlo Fortino. "A hybrid feature extraction method with regularized extreme learning machine for brain tumor classification." IEEE Access 7 (2019): 36266-36273.
- [16] Morad, Ameer Hussian, and Hadeel Moutaz Al-Dabbas. "Classification of brain tumor area for MRI images." In Journal of Physics: Conference Series, vol. 1660, no. 1, p. 012059. IOP Publishing, 2020.
- [17] NagaJyothi, Grande, and Sriadibhatla SriDevi. "Distributed arithmetic architectures for FIR filters-a comparative review." 2017 International conference on wireless communications, signal processing and networking (WiSPNET). IEEE, 2017.
- [18] Afshar, Parnian, Konstantinos N. Plataniotis, and Arash Mohammadi. "Capsule networks for brain tumor classification based on MRI images and coarse tumor boundaries." In ICASSP 2019-2019 IEEE international conference on acoustics, speech and signal processing (ICASSP), pp. 1368-1372. IEEE, 2019.
- [19] Das, Sunanda, OFM Riaz Rahman Aranya, and Nishat Nayla Labiba. "Brain tumor classification using convolutional neural network." In 2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT), pp. 1-5. IEEE, 2019.
- [20] Karthick Raghunath, K. M., Manjula Sanjay Koti, R. Sivakami, V. Vinoth Kumar, Grande NagaJyothi, and V. Muthukumar. "Utilization of IoT-assisted computational strategies in wireless sensor networks for smart infrastructure management." International Journal of System Assurance Engineering and Management (2022): 1-7.
- [21] Salçin, Kerem. "Detection and classification of brain tumours from MRI images using faster R-CNN." Tehnički glasnik 13, no. 4 (2019): 337-342.
- [22] Ari, Ali, and Davut Hanbay. "Deep learning based brain tumor classification and detection system." Turkish Journal of Electrical Engineering and Computer Sciences 26, no. 5 (2018): 2275-2286.
- [23] David, D. Stalin, and A. Jayachandran. "Robust classification of brain tumor in MRI images using salient structure descriptor and RBF kernel-SVM." TAGA Journal of Graphic Technology 14, no. 64 (2018): 718-737.
- [24] Vinoth Kumar, V., Karthikeyan, T., Praveen Sundar, P. V., Magesh, G., & Balajee, J. M. "Aspect based sentiment analysis and smart classification in uncertain feedback pool." International Journal of System Assurance Engineering and Management (2021): 1-11.
- [25] Seetha, J., and S. Selvakumar Raja. "Brain tumor classification using convolutional neural networks." Biomedical and Pharmacology Journal 11, no. 3 (2018): 1457.
- [26] Kumar, V. V., Raghunath, K. K., Rajesh, N., Venkatesan, M., Joseph, R. B., & Thillaiarasu, N. "Paddy plant disease recognition, risk analysis, and classification using deep convolution neuro-fuzzy network." Journal of Mobile Multimedia (2022): 325-348.
- [27] Abd-Ellah, Mahmoud Khaled, Ali Ismail Awad, Ashraf AM Khalaf, and Hesham FA Hamed. "Design and implementation of a computer-aided diagnosis system for brain tumor classification." In 2016 28th International Conference on Microelectronics (ICM), pp. 73-76. IEEE, 2016.
- [28] Belaid, Ouiza Nait, and Malik Loudini. "Classification of brain tumor by combination of pre-trained vgg16 cnn." Journal of Information Technology Management 12, no. 2 (2020): 13-25.
- [29] Bhanothu, Yakub, Anandhanarayanan Kamalakannan, and Govindaraj Rajamanickam. "Detection and classification of brain tumor in MRI images using deep convolutional network." In 2020 6th international conference on advanced computing and communication systems (ICACCS), pp. 248-252. IEEE, 2020.
- [30] Maithili, K., V. Vinothkumar, and P. Latha. "Analyzing the security mechanisms to prevent unauthorized access in cloud and network security." Journal of Computational and Theoretical Nanoscience 15.6-7 (2018): 2059-2063.
- [31] Zacharaki, Evangelia I., Sumei Wang, Sanjeev Chawla, Dong Soo Yoo, Ronald Wolf, Elias R. Melhem, and Christos Davatzikos. "MRI-based classification of brain tumor type and grade using SVM-RFE." In 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, pp. 1035-1038. IEEE, 2009.

*Edited by:* Polinpapilinho Katina

*Special issue on:* Scalable Dew Computing for Future Generation IoT systems

*Received:* Jul 28, 2023

*Accepted:* Nov 15, 2023



## A CLASS SPECIFIC FEATURE SELECTION METHOD FOR IMPROVING THE PERFORMANCE OF TEXT CLASSIFICATION

VENKATESH V\*, SHARAN S B†, MAHALAXMY S‡, S MONISHA§, ASHICK SANJEY D S¶ AND ASHOKKUMAR P||

**Abstract.** Recently, a significant amount of research work has been carried out in the field of feature selection. Although these methods help to increase the accuracy of the machine learning classification, the selected subset of features considers all the classes and may not select recommendable features for a particular class. The main goal of our paper is to propose a new class-specific feature selection algorithm that is capable of selecting an appropriate subset of features for each class. In this regard, we first perform class binarization and then select the best features for each class. During the feature selection process, we deal with class imbalance problems and redundancy elimination. The Weighted Average Voting Ensemble method is used for the final classification. Finally, we carry out experiments to compare our proposed feature selection approach with the existing popular feature selection methods. The results prove that our feature selection method outperforms the existing methods with an accuracy of more than 37%.

**Key words:** Feature selection, machine learning, class-specific feature selection.

**1. Introduction.** The textual data volume is increasing day by day. This causes the need for a text classification that categorizes the text into two or more classes. Increasing text classification performance involves many stages such as feature selection, feature weighting, instance selection, and instance weighting [1]. Feature selection is the process of picking a subset of features that maximizes the performance of a machine-learning model. Feature weighting is a standard to assign a score to each feature based on its importance. Instance selection involves identifying proper instances that help the machine learning model to learn the relationship between the features and class much more easily. Similar to feature weighting, instance weighting assigns an importance score to each instance. This paper focuses on feature selection.

Feature selection is one of the important methods which is applied to the text classification problem. In feature selection, a subset of features is selected which can be used in the classification models which outputs an increased accuracy [2]. Much research shows that only a handful of features are enough to build a good classification model rather than the entire set of features. There are lots of advantages when the feature selection is applied which include dimension reduction, increased classification performance, removal of redundancy, and so on [3].

Feature selection aims to remove the useless features and retain only the precise features that help the classifier to understand the relationship between the features and classes [4]. The feature selection methods can be categorized into filter types, wrapper types, and embedded types. The filter-based feature selection method relies to statistical information and selects the best subset of features. Whereas the wrapper method leverages the performance of feature classification by selecting the optimal subset of features. Unlike filter methods, the wrapper methods are classifier-dependent, which means if the classification model changes, then the subset

---

\*Department of Cybersecurity and Internet of Things, Sri Ramachandra Faculty of Engineering and Technology, Sri Ramachandra Institute of Higher Education and Research

†Department of Cybersecurity and Internet of Things, Sri Ramachandra Faculty of Engineering and Technology, Sri Ramachandra Institute of Higher Education and Research

‡Department of Artificial Intelligence and Data Analytics, Sri Ramachandra Faculty of Engineering and Technology, Sri Ramachandra Institute of Higher Education and Research

§Department of Cybersecurity and Internet of Things, Sri Ramachandra Faculty of Engineering and Technology, Sri Ramachandra Institute of Higher Education and Research

¶Department of Cybersecurity and Internet of Things, Sri Ramachandra Faculty of Engineering and Technology, Sri Ramachandra Institute of Higher Education and Research

||Department of Artificial Intelligence and Machine Learning, Sri Ramachandra Faculty of Engineering and Technology, Sri Ramachandra Institute of Higher Education and Research

of selected features is also changed. Embedded methods are the mix of filter and wrapper methods [5]. The feature selection method used in this manuscript is based on filter-based feature selection.

Filter methods are proven to be must faster and provide more advantages over the rest of the two types [6]. As in text classification, the number of features is very high thus creating a high dimensional space with many zeros in the feature matrix. The filter-based methods help to reduce the dimensions and make the classification model more efficient. Moreover, the filter-based methods are powerful in removing inconsistencies such as noisy data, and redundancies between features [7]. In this paper, we make use of filter-based feature selection to extract a promising feature subset.

Feature selection can be general or class-specific. In the first case, the feature selection algorithm picks a set of features that represents the entire class as a whole, whereas the second type of feature selection helps to pick features that are specific to each class [8]. The selected features may not be the same for all the classes, hence the advantage of class-specific feature selection is, that the model can learn the unique aspects of a class.

Based on the literature, the following limitations are found in the text classification domain.

- The existing feature selection methods do not capture unique features to represent a class uniquely
- The imbalance dataset classification remains a challenge in the field of text classification
- The problem of merging multiple binary classifications into a single multi-class classification needs to be addressed.

We propose a novel class-specific feature selection method with multiple levels of classification to improve the text classification performance.

In this manuscript, we develop a unique class-specific feature selection method that helps the machine learning model to understand the relationship between the input and output variables much deeper. The use of class-specific feature selection comes with its challenges such as imbalanced instances during the training phase, difficulty in ensemble the result of classification, and so on [10] [9]. We focus on optimizing the class-specific feature selection methods in text classification by the use of sampling and ensemble voting. The main contributions of the paper are as follows.

1. To propose a text classification model that has strong knowledge about all the classes.
2. To eliminate the class imbalance problem by using sampling and similarity-based elimination.
3. To increase the performance of the machine learning classification by introducing ensembling methods.

The remainder of this paper is organized in the following manner. Some recent related works regarding the class-specific feature selection are reviewed in section 2. The section 3 explains the problem statement of this manuscript. Section 4 elaborates the proposed classification model on class-specific features. Section 5 compares the proposed approach with the popular machine learning models by using a standard benchmark dataset. Section 6 ends this paper by providing the conclusion and scheduling a few future works.

**2. Related Works.** A four-stage framework which includes binarization, balancing, feature extraction, and finally, classification was used by [11]. The authors have used four machine learning models namely kNN (k Nearest Neighbour), SVM (Support Vector Machine), NB (Naive Bayes), and XGBoost for classification. Furthermore, eight feature selection methods have been used in their experiment. The dataset is split into training and testing by k fold method where k varies between 1 and 10. The results show that when k=10, all the model produces good performance with the maximum accuracy of 98.84 %.

A deep learning framework was proposed by [12]. In their work, the medical image classification is done using class-specific feature extraction methods. The final image classification is done by combining multiple convolution encoder-decoder frameworks. All the mini models learn about a discriminative, non-redundant training set. An 86.73 % accuracy was obtained by their proposed work.

A feature weighting method introduced by [13] uses the NB model for the classification as it is easy to use in text domains because the features are independent of each other. The authors suggested considering four conditions to measure the degree of relevance. The first one is the frequency of the term in the training dataset. The second one is the frequency of the term in the specific class. The third one represents the distribution of the class which contains the term and finally, the number of classes that contain the term. An accuracy of 86.05 % was obtained using their proposed algorithm.

A work by [14] uses class-specific feature extraction and principal component analysis for classification. They extend the traditional dictionary learning for all classes to specific classes. Each class has its internal

dictionary to express more about itself. Eight datasets are used to validate their proposed work.

In some cases, the features of a dataset cannot be used to classify the instances into one or more classes, in that case, the dataset is termed an uncertain dataset. To classify an uncertain dataset, rough-based features are used. The authors in [15] use the kNN model to extract class-specific features from an uncertain dataset using rough set theory.

Building a classification model that works with high-dimensional data is very challenging. A work by [16] uses the J48 classifier to reduce the dimensions of the data and then extract the class-specific features for each class using the term weighting approach. The authors show that the method runs much faster than the traditional methods.

A work proposed by [17] extracts class-specific features that automatically determine the number, length, and start values. This work reduces the number of calculations required by the traditional models and thus increases the speed of calculations. Since the features are extracted for each class separately, all these processes can be done parallelly to reduce the time.

The authors in [18] present a class-specific feature extraction model which samples a small number of training patterns from the original data, then for each class in the dataset, the algorithm extracts the patterns. The authors introduce a new classification algorithm that uses all the extracted class features to determine the classes for the instances.

Another interesting work by [19] uses three metrics to extract class-specific feature subsets. The first metric is the correlation between the features and the target class. The next metric is the relevance of the features in the selected subset and the last metric is the redundancy of the features in the selected subset.

**3. Problem Statement.** Let the text corpus  $D$  represented by  $N$  number of documents as shown in Eq 3.1. The goal is to classify each of the  $N$  documents to one of the classes  $c_i$  where  $c_i \in C$

$$D = \{d_1, d_2, \dots, d_N\} \quad (3.1)$$

Each unique document in  $D$  is represented as a tuple of features and classes as presented in Eq 3.2.

$$D = (x_i, c)^M \quad (3.2)$$

The goal is to perform class binarization which creates unique sub-sets of features for each class which can be expressed as Eq 3.3.

$$FS_i = \begin{cases} (x, 1) \forall x \in X & , if y = i \\ (x, 0) \forall x \in X & , if y \neq i \end{cases} \quad (3.3)$$

$$F_c = argmax_f I(f, c) \quad (3.4)$$

After the feature sets are extracted, the subset is used to train the machine learning models and predict the outcome as shown in Figure in 3.1. The candidate features are identified for each class, these features represent the class's unique features which can be the best for binary tests as denoted by Eq 3.4. To increase the efficiency, three machine learning model is used in each binary classification, and the weighted ensemble method is used to merge the result. The final classification is done by performing a mode ensemble approach.

**4. The Class Specific Feature Extraction Approach.** Feature extraction is a proven method to optimize the classification accuracy of a machine learning model, however, when the feature extraction is done on the entire dataset, the extracted subset of features may not represent the uniqueness of any one or more classes. This problem is overcome by a class-specific feature selection algorithm, where the features are extracted for each class and the subsets of two different classes may not be the same.

The first step is class binarization where we convert an  $N$ -class classification problem into an  $N$  two-class classification model (that is, we have only two classes- the positive and the negative). We start this process by splitting all the instances into two categories for each class, in one category, all the instances of the respective class are placed. In the other category, the rest of the entire instances are placed as shown in Figure 4.1. The main problem that is faced in this method is that, in each classification category, a huge imbalanced dataset is created. To eliminate the imbalance problem, we provide the following two solutions.

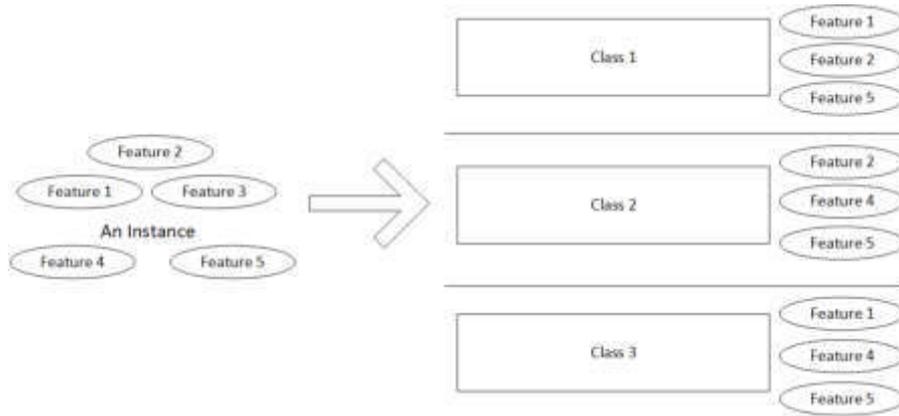


Fig. 3.1: The class-specific feature extraction process

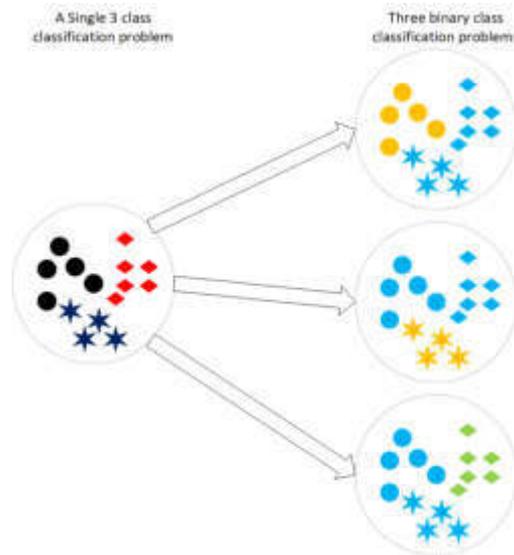


Fig. 4.1: The process of class binarization

1. Delete the similar instances in the majority class (the negative class)
2. Create instances using SMOTE (Synthetic Minority Oversampling Technique)

**4.1. Feature Selection.** Feature Selection plays an important role in data classification as it can shrink the dimension in the feature space and remove the redundancies. In this paper, we make use of three feature selection methods namely Information Gain (Eq 4.1), Gini Index (Eq 4.2), and Chi-Square (Eq 4.3). Information Gain is a popular feature selection method that measures how much information is provided by a feature to determine the target class. The higher the information, the more powerful the feature is. Once the information gained for all the features is found, then the top k features are selected and used to train the machine learning model.

$$H(x) = - \sum_i P_X(X_i) \log(P_X(X_i)) \tag{4.1}$$

$$Gini = \sum_{i=1}^m P(t | c_i)^2 P(c_i | t)^2 \quad (4.2)$$

$$Chi = \frac{N * (tp * tn - fp * fn)}{(tp + fp) + (tp + fn) + (tn + fp) + (tn + fn)} \quad (4.3)$$

**4.2. Deletion of Majority Instances.** As the N number of classes is mapped into two classes of size 1 (positive) and N-1 classes (negative), the negative class size becomes very large. This creates a huge imbalanced dataset, thus making the classification model learn more about one class and less about another class. To create a balanced dataset, one approach is to delete the instance from the majority instance (the negative class). The deletion of instances should be done with proper care so that the important instance should not be deleted and also the exact relationship should be learned from the dataset. To implement the deletion, first, the most similar instances have to be identified in the feature space. Then just by keeping one of the similar instances, all other instances can be deleted. The main advantage of the above-mentioned method is that random deletion is prevented.

Similar instances can be grouped by using any distance matching method that satisfies the following properties (assume  $\Theta$  is the distance finding algorithm).

1.  $\Theta(X, Y)$  is always equal to  $\Theta(Y, X)$
2.  $\Theta(X, X)$  is always 0.
3.  $\Theta(X, Y)$  is always  $> 0$ .

The similarity finding equation is presented at Eq 4.4. The function returns all the instances whose distance is less than a small threshold value  $\Delta$ .

$$Similar(X) = \{\forall x, x \in U, \Theta(X, x) \leq \Delta\} \quad (4.4)$$

Once similar instances for each instance are found, we have many sets of instances. In each set, the instances inside them reflect the same spatial properties in the feature space, so it is safe to delete the entire members of the set by keeping only one. This ensures that the model can learn all the relationships between the input features and the target classes in the dataset.

---

#### Algorithm 1 Local Classification

---

**Input:** Instances with all features

**Output:** The class associated with it

$F_1 \leftarrow$  Features selected by eq 4.1

$F_2 \leftarrow$  Features selected by eq 4.2

$F_3 \leftarrow$  Features selected by eq 4.3

Let  $M_1, M_2$  and  $M_3$  be three classification models build on  $F_1, F_2$ , and  $F_3$  respectively

$y = W_{F_1} * M_{F_1} + W_{F_2} * M_{F_2} + W_{F_3} * M_{F_3}$

return y

---

**4.3. Instance Generation using SMOTE.** In this manuscript, we use two methods to bring an imbalanced dataset into a balanced dataset, the first one is to delete similar instances (of the majority class) in the feature space. The other method is to generate synthetic instances. The most popular method used in generating synthetic instances is SMOTE (Synthetic Minority Oversampling) [20]. Many problems like overfitting, and loss of important information are addressed in SMOTE. In recent years, many optimal versions of SMOTE are available [21].

One of the main advantages of the SMOTE is, that it does not generate random instances which creates an information loss in the training phase of the classification model. The SMOTE picks a random sample in the minority class and finds the k closest neighbors using the Euclidean distance finding formula. Then a new location is identified within the selected instances from the Euclidean formula and a new synthetic instance is created.

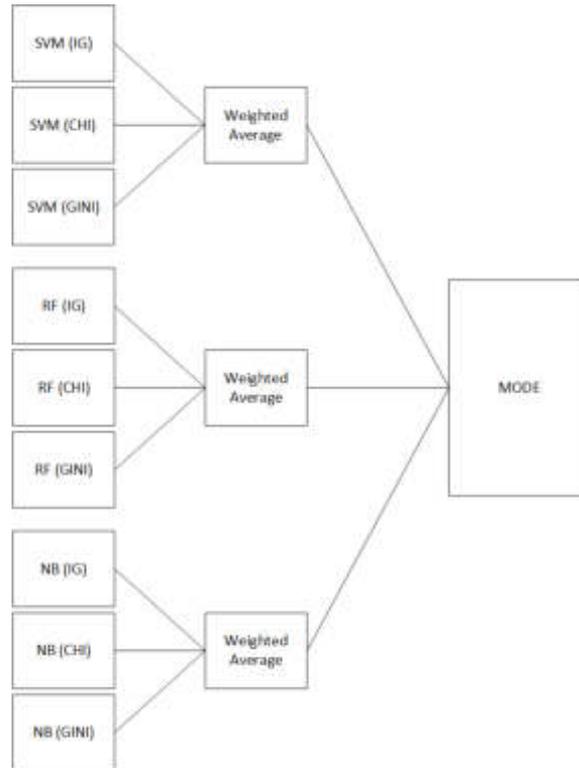


Fig. 4.2: The working of the classification model

**4.4. Machine Learning Based Classification.** We have used three machine learning models namely SVM, RF, and NB to classify the text documents and assign a class to them. We have adapted ensemble methods at multiple levels to increase the performance of the classification. We have selected these three machine learning models because of some reasons which we have explained in the following subsections.

**4.4.1. Support Vector Machines.** The SVM classifier can classify both linear and non-linear data. The SVM creates a non-linear hyperplane that can separate the instances. The main challenge an SVM model faces is to place the hyperplane in a perfect position that acts as a decision boundary that partitions the data instances. In text classification, the SVM works better as the feature space has many dimensions [22].

**4.4.2. Naive Bayes.** Naive Bayes is one of the statistical-based machine learning models that operate on the Bayes theorem. The semantics and underlying relationship between the words can be easily captured by using Naive Bayes. Since the classifier depends on conditional probabilities, it can be suited for a large dataset and gives optimal results.

**4.4.3. Random Forest.** Random Forest is one of the ensemble classifiers which combines multiple decision trees. Since multiple decision trees are used, each tree can be used to work on a feature. This ensures that the irrelevant features are not passed to other trees. The power of RF is the ability to generalize multiple trees.

**4.4.4. Multi Level Ensemble Methods.** For each class, we have each of these three machine learning models deployed three times one for each feature selection method. The classification results are finalized for each machine learning model based on the weighted average method as described in Eq 4.5. The full working of the ensemble model is described in algorithm 1. In the algorithm, the unique features for all the classes are extracted using three feature selection methods, and then the three classifiers such as SVM, NB, and RF are used to calculate the target class of the instance. The weighted voting ensemble method is used to find the

Table 5.1: Dataset Description

Sno	Dataset Name	# of classes	# of documents
1	Reuters-21578	10	9980
2	IMDB	2	50K
3	20newsgroup	20	18K

Table 5.2: Hyper parameters of ML model

Machine Learning Model	Hyperparameters
SVM	C=100, kernel = rbf
RF	estimators = 50

final class.

$$y = W_{IG} * M_{IG} + W_{CHI} * M_{CHI} + W_{GINI} * M_{GINI} \quad (4.5)$$

After each of the weighting average results are obtained the results are ensembled using majority ensembling as shown in Eq 4.6.  $y_1$ ,  $y_2$ , and  $y_3$  represent the outputs of the three weighted average methods. The abovementioned process is repeated parallelly for all the classes and the final class is determined by one hot ensembling. The working of the proposed approach is displayed in Figure 4.2.

$$y = mode(y_1, y_2, y_3) \quad (4.6)$$

**5. Results and Discussion.** We have experimented to test our proposed approach with three datasets. The details of the dataset are exhibited in Table 5.1.

In this experiment, we have used 10-fold validation. In this method, the dataset is randomly divided into 10 subsets. The proposed classifier is trained and tested 10 times where in each iteration, one subset is used as testing, and the remaining 9 subsets are used as training. The experiments are carried out in an i9 CPU with 16GB RAM. The details of hyper-parameters used in the experiment are shown in Table 5.2.

All the documents in the three datasets are pre-processed in the following order. First, the stop words are removed from the documents. Second, each word is stemmed and finally, all the features are represented in the bag of words model.

To evaluate the proposed approach, we have used four metrics known as accuracy, precision, and, recall. The metrics are calculated by using the formulas mentioned in Eq 5.1, 5.2 and, 5.3 respectively.

$$ACC = (TP + TN)/(TP + TN + FP + FN) \quad (5.1)$$

$$P = (TP)/(TP + FP) \quad (5.2)$$

$$R = (TP)/(TP + FN) \quad (5.3)$$

Table 5.3 provides the performance evaluation of all the machine learning models along with our proposed approach. It can be noted that the proposed approach outperforms all the existing classification models in the comparison. In existing models, NB produces the highest accuracy and recall whereas the SVM produces the highest precision.

From Table 5.4, it can be found that among all the classifiers in the comparison, the proposed approach exhibits the highest performance and has an advantage over the other machine learning models in the IMDB dataset. In the existing models, NB yields the highest performance in terms of accuracy, precision, and recall.

Table 5.3: Performance Evaluation of 20newspaper dataset

Model	Acc	Prec	Rec
kNN	0.75	0.8	0.75
RF	0.76	0.74	0.75
SVM	0.76	0.81	0.75
NB	0.77	0.80	0.77
LR	0.54	0.55	0.54
<b>Proposed</b>	<b>0.83</b>	<b>0.86</b>	<b>0.77</b>

Table 5.4: Performance Evaluation of IMDB

Model	Acc	Prec	Rec
kNN	0.62	0.75	0.67
RF	0.7	0.71	0.71
SVM	0.75	0.8	0.79
NB	0.76	0.81	0.75
LR	0.61	0.65	0.64
<b>Proposed</b>	<b>0.85</b>	<b>0.79</b>	<b>0.77</b>

Table 5.5: Performance Evaluation of Reuters Dataset

Model	Acc	Prec	Rec
kNN	0.81	0.81	0.8
RF	0.81	0.81	0.81
SVM	0.84	0.85	0.84
NB	0.85	0.85	0.85
LR	0.79	0.78	0.8
<b>Proposed</b>	<b>0.9</b>	<b>0.9</b>	<b>0.89</b>

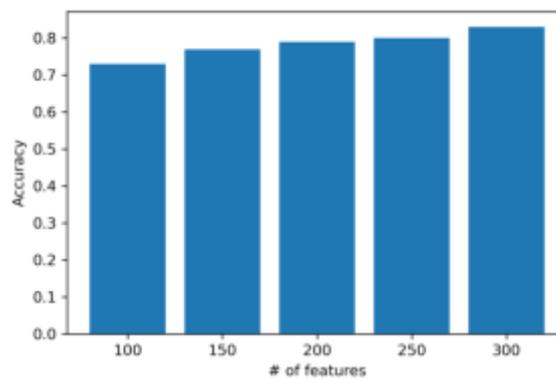


Fig. 5.1: The accuracy comparison of our proposed approach with respect to various number of features - 20newsgroup

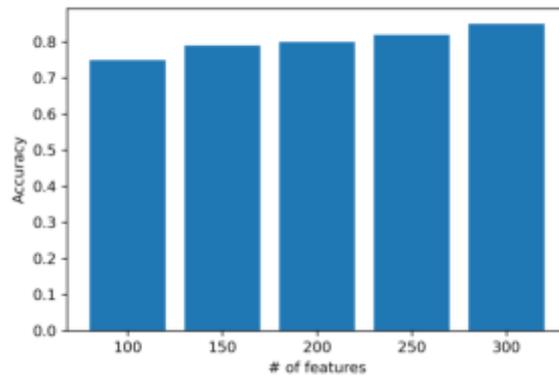


Fig. 5.2: The accuracy comparison of our proposed approach with respect to various number of features - IMDB

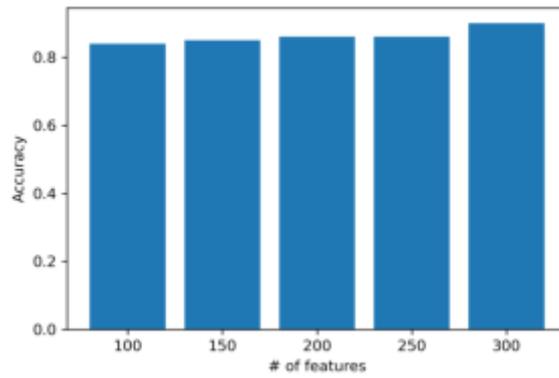


Fig. 5.3: The accuracy comparison of our proposed approach with respect to various number of features - Reuters-21578

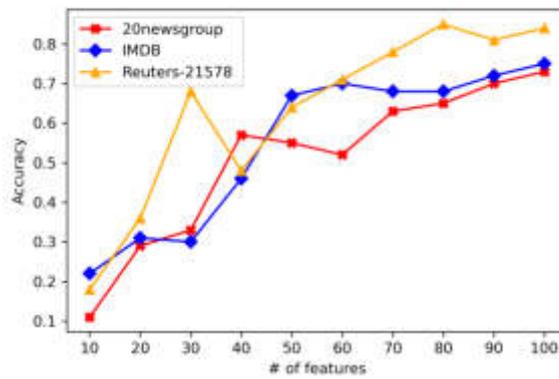


Fig. 5.4: The accuracy comparison of our proposed approach with respect to various number of features

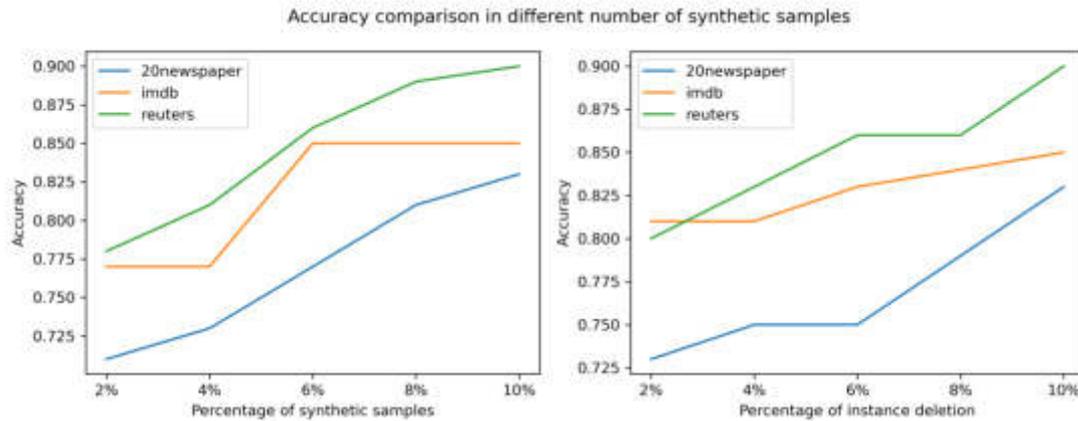


Fig. 5.5: The Accuracy Comparison in Imbalanced Dataset Tuning

The performance of the Reuters dataset is displayed in Table 5.5. The proposed classification model gives the best comparison result to the existing machine learning models. Naive Bayes produces the next highest accuracy and precision. The SVM produces the next highest recall.

From the Figures 5.1, 5.2 and 5.3, it can be observed that when the number of selected features increases, there is an increase in the performance of the machine learning model. When the number of selected features is fixed at 100, the relationship between the input features and the target class is not fully captured, and that's the reason for the low performance. Meanwhile, when the selected number of features increases, the machine learning model gets more information to learn about the relationship between the features and the target class and hence yields more accuracy. When the number of selected features is very low, the proposed approach is not producing satisfactory results. Figure 5.4 displays the accuracy when the number of selected features is less number of features. Moreover, we found that when the number of selected features exceeds 300, the classification performance for all three datasets becomes stable and no significant changes are found.

It is noted that the text classifiers perform much better in a balanced dataset than in an imbalanced dataset. There are two optimizations done to convert the imbalanced dataset into a balanced dataset. The first one is the creation of synthetic instances. Proper care is taken so that the created sample does not act as the original instance. The second one is the deletion of similar instances in the majority class without altering the original distribution of the dataset. Figure 5.5 displays the accuracy comparison in the two optimizations. The optimal value of both is 10%.

**6. Conclusion and Future Work.** Feature selection is one of the promising methods to increase the performance of the classification model. However, when the feature selection algorithm considers the entire class, then the selected features may not represent the unique characteristics of the classes. To solve this problem, class-specific features are selected so each class has its own set of feature subsets. This ensures that each subset fully represents the class and can hold the unique properties of the class. We propose a classification model that works on class-specific features and introduces ensembles on multiple levels. The experiment results prove that the proposed approach is effective and superior to the existing machine learning models. As the results are promising, the future scope of this research is planned to extend the work by incorporating deep learning approaches.

## REFERENCES

- [1] A. Palanivinaayagam, C. Z. El-Bayeh, and R. Damaševičius, "Twenty years of machine-learning-based text classification: A systematic review," *Algorithms*, vol. 16, no. 5, 2023.
- [2] J. Maldonado, M. C. Riff, and B. Neveu, "A review of recent approaches on wrapper feature selection for intrusion detection," *Expert Systems with Applications*, vol. 198, p. 116822, 2022.

- [3] U. M. Khaire and R. Dhanalakshmi, "Stability of feature selection algorithm: A review," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 4, pp. 1060–1073, 2022.
- [4] X. Liu, G. Zhou, M. Kong, Z. Yin, X. Li, L. Yin, and W. Zheng, "Developing multi-labeled corpus of Twitter short texts: A semi-automatic method," *Systems*, vol. 11, no. 8, 2023.
- [5] S. Bazzaz Abkenar, M. Haghi Kashani, M. Akbari, and E. Mahdipour, "Learning textual features for Twitter spam detection: A systematic literature review," *Expert Systems with Applications*, vol. 228, p. 120366, 2023.
- [6] L. Hu, L. Gao, Y. Li, P. Zhang, and W. Gao, "Feature-specific mutual information variation for multi-label feature selection," *Information Sciences*, vol. 593, pp. 449–471, 2022.
- [7] G. S. Thejas, R. Garg, S. S. Iyengar, N. R. Sumitha, P. Badrinath, and S. Chennupati, "Metric and accuracy ranked feature inclusion: Hybrids of filter and wrapper feature selection approaches," *IEEE Access*, vol. 9, pp. 128687–128701, 2021.
- [8] U. L. Altintakan and A. Yazici, "Towards effective image classification using class-specific codebooks and distinctive local features," *IEEE Transactions on Multimedia*, vol. 17, no. 3, pp. 323–332, 2015.
- [9] C. P. Ezenkwu, U. I. Akpan, and B. U.-A. Stephen, "A class-specific metaheuristic technique for explainable relevant feature selection," *Machine Learning with Applications*, vol. 6, p. 100142, 2021.
- [10] K. Liu, T. Li, X. Yang, X. Yang, and D. Liu, "Neighborhood rough set based ensemble feature selection with cross-class sample granulation," *Applied Soft Computing*, vol. 131, p. 109747, 2022.
- [11] D. K. Rakesh and P. K. Jana, "A general framework for class label specific mutual information feature selection method," *IEEE Transactions on Information Theory*, vol. 68, no. 12, pp. 7996–8014, 2022.
- [12] U. Upadhyay and B. Banerjee, "Compact representation learning using class specific convolution coders - application to medical image classification," in *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, pp. 1266–1270, 2020.
- [13] S. Ruan, H. Li, C. Li, and K. Song, "Class-specific deep feature weighting for naïve bayes text classifiers," *IEEE Access*, vol. 8, pp. 20151–20159, 2020.
- [14] F. Pan, Z.-X. Zhang, B.-D. Liu, and J.-J. Xie, "Class-specific sparse principal component analysis for visual classification," *IEEE Access*, vol. 8, pp. 110033–110047, 2020.
- [15] M. Sewwandi, Y. Li, and J. Zhang, "A class-specific feature selection and classification approach using neighborhood rough set and k-nearest neighbor theories," *Applied Soft Computing*, vol. 143, p. 110366, 2023.
- [16] R. Patil and V. M. Barkade, "Class-specific features using j48 classifier for text classification," in *2018 Fourth International Conference on Computing Communication Control and Automation (ICCCUBEA)*, pp. 1–5, 2018.
- [17] Z. Liang and H. Wang, "Efficient class-specific shapelets learning for interpretable time series classification," *Information Sciences*, vol. 570, pp. 428–450, 2021.
- [18] A. Roy, P. D. Mackin, and S. Mukhopadhyay, "Methods for pattern selection, class-specific feature selection and classification for automated learning," *Neural Networks*, vol. 41, pp. 113–129, 2013. Special Issue on Autonomous Learning.
- [19] X.-A. Ma, H. Xu, and C. Ju, "Class-specific feature selection via maximal dynamic correlation change and minimal redundancy," *Expert Systems with Applications*, vol. 229, p. 120455, 2023.
- [20] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "Smote: Synthetic minority over-sampling technique," *Journal of Artificial Intelligence Research*, vol. 16, p. 321–357, 2002.
- [21] A. Zhang, H. Yu, S. Zhou, Z. Huan, and X. Yang, "Instance weighted smote by indirectly exploring the data distribution," *Knowledge-Based Systems*, vol. 249, p. 108919, 2022.
- [22] A. Palanivinaiyagam and R. Damaševičius, "Effective handling of missing values in datasets for classification using machine learning methods," *Information*, vol. 14, no. 2, 2023.

*Edited by:* Polinpapilinho Katina

*Special issue on:* Scalable Dew Computing for Future Generation IoT systems

*Received:* Aug 10, 2023

*Accepted:* Sep 20, 2023



## BREAST CANCER IMAGE CLASSIFICATION BASED ON ADAPTIVE INTERPOLATION APPROACH USING CLINICAL DATASET

SUSMITHA UDDARAJU\* A KOUSHYLA† I HEMALATHA‡ M MARAGATHARAJAN§ BALASUBRAMANIAN C¶ AND L SATHISH KUMAR||

**Abstract.** In the healthcare and bioinformatics disciplines, the categorization of breast cancer has become an emerging paradigm due to the second most common cause of cancer-related mortality in women. A biopsy is a procedure in which tissue has been examined to determine whether or not it is breast cancer through histopathologists that may lead to a mistaken diagnosis. The research is mainly focused on the patient data (884 case reports of patients) is acquired from the American Oncology Institute implemented with preprocessing techniques consists of missing values and those values are recovered with Novel Modified Interpolation (MI) Method. Deep learning networks effectively detect and assess the pattern for annotating histological data based on the labelling, which preserves time, system cost and enhances the system accuracy. This framework addressed feature acquisition and missing analysis strategies based on entropy confidence weight factor. First, the iterative patterns have been treated as a potential diagnostic rule, and the attention-based rule combination formulates the classification issue based on integrating convolutional and recurrent neural networks, and the short-term and long-term spatial correlations between patches. Second, the key part of label construction is carried out with an entropy confidence-weight factor assessment which detects and predict different patterns to construct the rule for classification. Third, optimization of clustering data by assessing missing parameters based on mean square error and the concept of interpolation to reduce data loss by around 20% and enhance the system accuracy. Simulation results show that the proposed system achieves 91.3% accuracy to state-of-art approaches, potentially allied in clinical applications. Modified Interpolation (MI) method recovered missing values with least mean square error and less data loss of 0.0123 and 1.38% respectively. This method also compared with existing Linear Interpolation (LI) method which is able to recover least mean square error and data loss of 3.295 and 18.925% respectively. Comparatively the modified interpolation method recovered the missing values with less mean square error and less data loss.

**Key words:** Breast Cancer, attention-based rule combination, Data Classification, Breast Cancer Diagnose, Data Clustering and Recovery

**1. Introduction and examples.** Breast cancer is the second leading cause of death in women, behind lung cancer, and it is the most frequent and lethal invasive cancer in women. According to the WHO's International Agency for Research on Cancer (IARC), 8.2 million people died in 2012 due to cancer-related causes. By 2030, experts predict that the number of new cases will have risen to almost 27 million [1]. Based on apparent signs, particularly in past decades, doctors have been keeping track of lumps since they may develop into tumours. It is true since breast lumps often appear as outwardly apparent tumours in contrast to other forms of cancer inside the body. Early discovery and diagnosis were difficult since the disease was stigmatized by social stigma and shame. In addition to medical periodicals and novels, literature seldom mentioned breast cancers. The involvement of women in active promotion and awareness is essential to mitigate this impact. Chemotherapy is a medicinal treatment that uses strong chemicals to control the disease but that damage body fast-growing cells. Chemotherapy is the regularly utilized disease therapy since disease cells develop and multiply altogether quicker than the other body's cells. Chemotherapy meds arrive in different structures. Chemotherapy drugs can treat a great many malignancies, either alone or in blend. Chemotherapy is a powerful therapy for some kinds of disease, yet it likewise accompanies a gamble of unfriendly impacts.

\*Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur Dist., AP, India ([susmithaударaju@gmail.com](mailto:susmithaударaju@gmail.com)).

†Krishna Collage of Engineering and Technology, Coimbatore, Tamilnadu, India ([kousalyaa@skcet.ac.in](mailto:kousalyaa@skcet.ac.in)).

‡Department of Information Technology, S. R. K. R. Engineering College, Bhimavaram, India ([indukurihemalatha@gmail.com](mailto:indukurihemalatha@gmail.com)).

§VIT Bhopal University, Kothrikalan, Madhya Pradesh, India ([maragatharajanm@gmail.com](mailto:maragatharajanm@gmail.com)).

¶Kalasalingam Academy of Research and Education, Krishnankoil, Tamilnadu, India ([baluece@gmail.com](mailto:baluece@gmail.com)).

||VIT Bhopal University, Kothrikalan, Madhya Pradesh, India ([lsathishkumarsva@yahoo.in](mailto:lsathishkumarsva@yahoo.in)).

Every year on October 19, people around the world observe International Breast Cancer Day of Awareness to raise awareness of breast cancer. For women, breast cancer makes up close to 30% of tumors. The objectives of World Breast Cancer Day 2022 are to increase awareness and support women's access to prompt and efficient prevention, detection, and treatment. Total number of 87,090 female died from Breast Cancer in the year of 2018 alone in India which is also represented in Fig. 3.1. Breast Cancer accounted for about 24.5% of all cancer related deaths in women in India [2].

Almost about one in four deaths due to cancer, in ladies in our country, was due to Breast Cancer. Mortality Rate of Different Cancer Patients [3] Once the patient diagnosed with breast cancer, the anticipated progression and outcome of your ailment will be estimated by the doctor. The prognosis varies from person to person and depends on a no of variables, like age of the patient, region of the patient, grade, and extent of the cancer. As you become older, your risk of getting breast cancer rises. women between the ages of 65 and 74 are the 2 ones who are diagnosed with breast cancer most frequently. Reliable Source. The median age that women are given breast cancer diagnoses is 63 years old. Less than 2% of the ladies in our country who received a BC diagnosis between 2015 and 2019 were under the age of 35.

Some chemotherapy incidental effects are minor and controllable, while others can life-compromise. In recent years, histopathological image usage and analysis accuracy has been increased due to the deployment of deep learning approaches. However, there are certain issues with deep learning approaches as follows:

1. Size of image data and high time consumption to train the data.
2. Computationally expensive when feeding raw data.

Few recent works have achieved moderate accuracy by stochastic selection of patches from image [2] or resizing the images [3, 4], but both mechanisms are affected by abnormal information loss that could lead to misdiagnosis. Histopathological image classification (HIC) assess different types of lesions (such as benign and malignant tumours). In [2], a sparsity model has been designed to encode image patches based on a fusion of attention patch predictions. Segmentation of the Whole slide image (WSI) patch is essential during cancer image classification, and this segmentation would be into blocks and super-pixels based on the HIC approach. Consequently, the content-based histopathological image retrieval (CBHIR) method probes regions with similar content towards ROI. This mechanism provides comprehensive information to classify the images and helps pathologists. CBHIR provides comprehensive information for effective diagnosis based on historical data analysis. As per the studies, CBHIR has achieved similar accuracy compared to HIC [19]. The CNN extract the features of each patch where a support vector machine (SVM) is utilized to classify histopathological image [5]. Usually, the histopathological image classification process splits the image into smaller patches; the CNN based patch classification results are used to classify the cancer image. Histopathological image resolutions sometimes cause adverse effects that may impact the accuracy due to ignoring long-distance spatial dependency. Sometimes, the feature patch selection is not prosperous because of significant data loss that leads to insufficient data fusion. Subsequently, there are few research challenges with open-source data sets like ImageNet for object detection, BreakHis dataset [18] and Bioimaging2015 dataset [7]. However, their diversity is not guaranteed. In this regard, the designed method assesses the multilevel features based on the patches' long and short-term spatial correlations. Second, the label construction is carried out with an entropy confidence-weight factor assessment which detects and predict different patterns to construct the rule for classification. Third, optimization of clustering data by assessing missing parameters based on mean square error and the concept of interpolation to reduce data loss. Finally, the RNN is deployed to fuse the patch features for final image classification. The 4-class classification task would achieve an accuracy of 91.3% than state-of-the-art methods. The proposed strategy is quicker and effective than the extant strategies. The usage can be considered as an augmentation of a medical clinic arrangement to monitor patient status distantly. Our contributions are listed below:

1. Design a attention-based rule combination to construct potential diagnostic rule for classification issue formulation.
2. Design an entropy confidence-weight factor to detects and predict different patterns prior to classification.
3. Design an optimization strategy to optimize clustering data loss based on mean square error and interpolation to enhance the system accuracy.

The rest chapter is compiled as: Section 2 describes related work with respective to feature extraction and its issues. Section 3 describes proposed method and attention-based rule combination scheme for effective classification. Section 4 describes implantation and results analysis. Section 5 describes conclusion of the manuscript.

**2. Related Work.** Many scientific domains rely on spatial and biological interpolation of spatio temporal and socioeconomic data. Nearest neighbour (NN), inverse distance weighting (IDW) [8], and trend surface mapping (TS) [9] are some basic interpolation approaches. Geostatistical interpolation (kriging) [10] was introduced in the 1980s. Because kriging considers spatial correlation and quantifies interpolation error using the kriging standard deviation, this was a significant improvement.

Under unambiguous stationarity suspicions, In spatial information kriging is the Best Linear Unbiased Predictor (BLUP)[11]. It is also relatively adaptable, as numerous variations can handle specific situations like anisotropy, non-normality, and information included in covariates [12].

Kriging, on the other hand, has drawbacks. It can be computationally intensive, and it might be challenging to develop a solid geostatistical model that effectively fits all sorts of data [13]. It is also not well suited to adding the vast amounts of covariate data that are now available. One significant challenge is that defining a geostatistical model for data that cannot be easily transformed to normalcy is difficult. Indicator kriging [14] was created to address this problem; however, it is badly designed and not model-based (i.e., it doesn't utilize formal factual strategies inferred for an unequivocal and complete measurable model, see Diggle and Ribeiro). The Summed up Direct Geostatistical Model [15] is statistically solid, but it has limitations regarding the types of distributions it can handle, and it is also technically challenging. For example, it is not always clear how variables with a lot of them interact.

Despite the fact that yearly and month to month precipitation in arid regions can in any case have zero qualities and solid positive skewness, spatial addition utilizing kriging is less complicated than every day or hourly precipitation in light of the fact that transiently collected precipitation keeps an eye on the typical dissemination[6]. However, when mapping hourly or daily precipitation, geographic variability is more significant, the stationarity assumption is called into question, and the precipitation distribution becomes skewed and filled with zeros [16]. From the literature, it is clear that there is a need for an extensive database to do adequate research. That cluster's data must be omitted if even one parameter is missing. Usually, 20% of the data may be missed because of a single missing parameter in the database. We designed an algorithm to recover missing parameters with a small mean square error to prevent this problem. The designed method assesses the multi-level features based on the patches' long and short-term spatial correlations. Second, the label construction is carried out with an entropy confidence-weight factor assessment which detects and predict different patterns to construct the rule for classification [17]. Third, optimization of clustering data by assessing missing parameters based on mean square error and the concept of interpolation to reduce data loss. At long last, the RNN is sent to combine the fix highlights for definite picture order.

**3. Proposed system.** Fig. 3.2 represents the Steps to perform and apply MI method as below:

1. Clustering of database.
2. Define overlap clustering for different classes.
3. Locate the cluster with missing parameters.
4. Apply modified interpolation to recover missing parameter.

Firstly, the collected real time data is clustered because the collected data is independent, we have to consider the classes. So we cluster our data by using output feature. Secondly, have to consider overlap clustering because there are many different classes, by using that classes that any data is overlapping. Thirdly, location of cluster with missing parameter is done to apply MI method. Once the location of the cluster is founded the MI model is applied to recover missing value.

The clustering approach is formulated based on spatial correlations, feature and instance dimension in a matrix format. Let assume  $m \times n$  matrix A, and matrix B is sub-set of matrix A; where m is number of rows refers cases, n is number of columns refers features. The tumors have assigned scores to identify the lesion type for breast tumour diagnosis. In most cases, the score based analysis is not optimal; therefore, constant features have considered to accomplish the image classification. The mean-square-residue score (MSRS) measures each

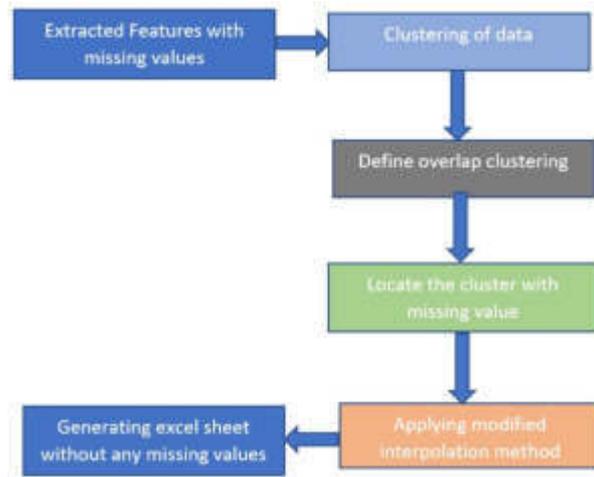


Fig. 3.1: Flowchart of recovering missing values.

Table 3.1: Tumor characteristics.

Characteristics of	benign	malignant
Shape	Oval	Irregular
Orientation	Parallel	Non-parallel
Border	distinct	indistinct
Margin	without spiculated	spiculated

cluster’s rate during the image classification. Each cluster enables at least 5 rows to formulate the unique rule in our simulation.

Initializes the required parameters, such as probability error and supportive confidence weight attributes. Step 1: check all image patches to construct the rule based on a cluster with an influential confidence factor. step-3 define the parameters, line-4 estimates the distance threshold value. step-6 to step-7 construct the cluster when the MSRS score is less than its threshold value. Otherwise, go to step-2, and this process is repeated till the construction of the cluster with an optimal confidence value. The final cluster remains used to construct the diagnosis rule, but there is a need to assess the confidence factor to validate the reliability. The rule construction must differentiate benign or malignant categories. Let us assume,  $K_{be}$  indicates benign-label rows,  $K_{ma}$  indicates malignant-label rows and  $K_{bm}$  indicates total rows of a cluster. The confidence-weight factor  $\phi$  is derived as follows

$$\phi_{be} = \frac{K_{be}}{K_{bm}} \tag{3.1}$$

$$\phi_{ma} = \frac{K_{ma}}{K_{bm}} \tag{3.2}$$

The essential confidence-weight factor  $\phi$  of image  $g$  or cluster is derived as follows

$$\phi_g = \max \{ \phi_{be}, \phi_{ma} \} \tag{3.3}$$

The cluster is classified based on the larger confidence-weight factor  $\phi$ . The cluster-based rule construction is conditionally formulated when the satisfaction of this condition  $\phi \geq \phi_{th}$ . In our simulation, the  $\phi_{th}$  is fixed with 0.75 to enhance the system accuracy.

**Algorithm 1** Pseudocode of spatial correlation method

---

```

Input: Collected data with missing values in the form of Excel
Read the excel sheet
Initialize an empty array called 'missingElements'
for each column in excel sheet:
if column name is 'No of Missing Values':
value of cell > 0:
add this row to 'missingElements' array
Initialize a data structure 'D'
for storing clusters for each element in 'missingElements' array:
pass the element to the Mathematical model:
use KNN clustering to form clusters consecutively  $C_1, C_2, C_3...$  and store in D
for each cluster  $C_i$  in D:
for each feature  $F_j$  in  $C_i$ :
for each value  $P_k$  in  $F_j$ :
 $\sum P$ : Sum of values  $P_k$  in the feature  $F_j$ 
PN: Number of missing values in feature  $F_j$ 
P*: missing value
Define Overlap Clustering for the above formed clusters and missing values.
Locate the cluster with missing values.
Apply Modified interpolation method to remove missing values.
Output: Generates new excel sheet by recovering missing value

```

---

Computer-aided pathology analysis relies on the ability to classify breast cancer data. The researcher proposes to use a modified interpolation technique with data clustering to identify missing characteristics in a breast cancer database. In this regard, first extracting useful and non-redundant characteristics that are not redundant.

Let us assume,  $Database \in \{C_1, C_2, \dots, C_N\}$  enables N set of clusters and each cluster is constructed based on set of features  $F_N$  such as  $C \in \{F_1, F_2, \dots, F_N\}$  and each feature is extracted based on set of parameters such as  $F \in \{P_1, P_2, \dots, P_N\}$ . Now the missing parameters are derived as follows

$$\hat{P} = \frac{\sum \frac{\partial P}{\partial F} + \sum P}{P_N} \quad (3.4)$$

where  $\frac{\partial P}{\partial F}$  is a derivative parameter of the concern feature. All possible combinations have been performed on the target variables. The nearest location is first calculated for the calibration. The map of the nearest calibration variables is then fed to the gradient calculation. The missing parameter equation is used to find out the nearest prediction parameters as shown in the proposed system architecture. During HIC, the object may have different scales and complexity that leads hard to characterize image features. In this regard, CNN has achieved high accuracy, but the HIC requires excellent attention to represent the features. Therefore, combining multi-layer features helps to achieve the targeted accuracy by retaining fine-grained details and local textures, as we can observe in Fig. 3.3. The Tensorflow slim distribution helps to assess the feature extraction. In our simulation, the fully connected layers have not been considered due to the arbitrary size of images. Each layer channel is combined to accomplish multilevel convolutional features based on global average pooling with two inception modules. The final vector with 4096D dimensions enable multi-features of the image to classify the breast cancer images.

**4. Experiment Result Analysis.** Using the Tensor Flow framework the proposed algorithm is executed. The CNN architecture is a basic deep model, an AlexNet is employed as pre-trained model on the ImageNet. The essential settings of the server incorporate an Intel 3.6-GHz CPU and NVIDIA GeForce GTX 1080Ti GPU. In this experiment, 1062 breast cancer instances ( 418 benign, 644 malignant cases) are considered. The SVM, fuzzy SVM, GRNN based methods are considered as state-of-art approaches. Fig. 4.1 illustrates different tissue

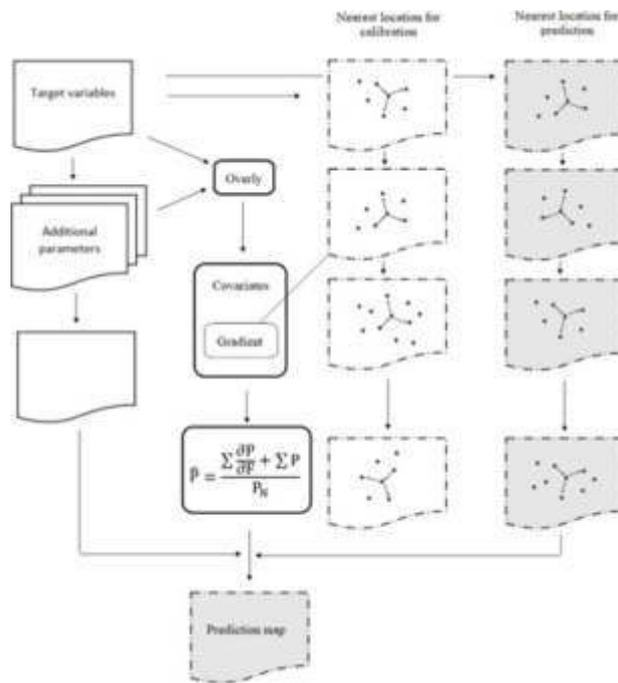


Fig. 3.2: Proposed System Architecture

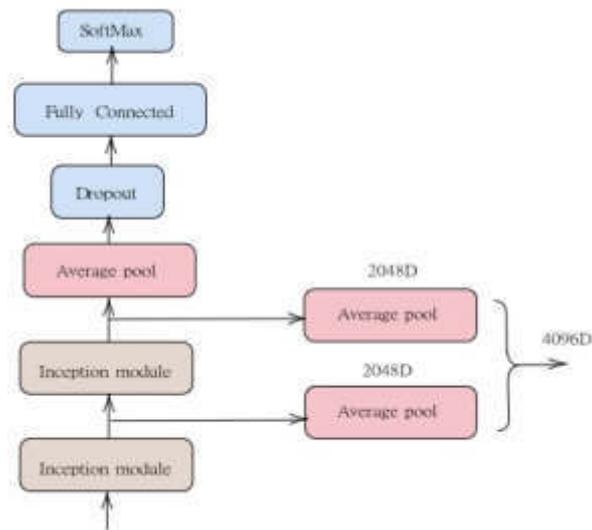


Fig. 3.3: Schematic overview of feature representation

samples which considered as classes to assess the performance of the system, and where a) is a normal tissue, b) is a benign tissue, c) is a In-situ carcinoma, d) is a Invasive carcinoma.

**4.1. Performance Metrics.** Distinctive execution measurements are used to decide the productivity of our approach. Well known system performance like Accuracy, sensitivity and specificity are measured. The

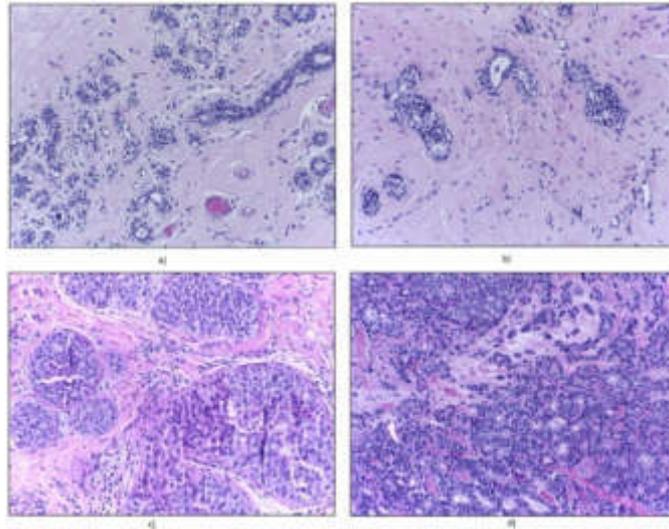


Fig. 4.1: Four different tissue samples

Table 4.1: Classification based on essential attributes.

S.No	ML Model	-1,1	-1,1	-1,1	TP	TN	FP	FN	Acc	TPR	TNR	FPR	FNR	PPV	NPV
1	SVM	138	54	28	138	536	54	28	89.15	83.13	90.85	9.15	16.87	71.88	95.04
2	GRNN	192	0	564	192	100	0	564	34.11	25.4	100	0	74.6	100	15.06
3	Fuzzy SVM	125	67	11	125	553	67	11	89.68	91.91	89.19	10.81	8.09	65.1	98.05
4	DL-AI approach	171	24	4	171	567	24	4	96.1	97.01	95.11	3.95	1.74	86.21	99.41

accuracy is the proportion of correct classification in all instances. The sensitivity is the proportion of positives (malignant instances) that are correctly identified, and the specificity is the proportion of negatives (benign instances) that are correctly identified, which are listed below.

$$Precision = \frac{(TP)}{(TP) + (FP)} \quad (4.1)$$

$$Accuracy = \frac{(TP) + (TN)}{((TP) + (TN)) + ((FP) + (FN))} \quad (4.2)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N \|P_i - Q_i\| \quad (4.3)$$

Exactness speaks average capacity of DRL model. True positive (TP) and true negative (TN) measure concludes the nonappearance and nearness of C-19 virus. False positive (FP) and false negative (FN) recognize the quantity of bogus caused by the system. Function measure (FM) remain used to decide the expectation execution rate. Root mean square error (RMSE), mean absolute error (MAE) have used to assess distinction rate among real and anticipated esteems. Table 4.1 illustrates that our DL-AI approach has achieved reasonable accuracy than other extant approaches. The precision rate is stable and DL-AI approach accomplished tremendous precision than the extant models because the designed Spatial Co-relation system effectively classify the images based on DRL system.

Table 4.2: Comparative analysis of proposed method with other models.

Study	Feature Extraction	Classifier Method	Classes	Accuracy (%)
1	WT	SVM	3	89.1
2	Morphology	GRNN	5	88.5
3	Floating Features	Fuzzy SVM	5	90
4	Spatial Co-relation	DL-AI approach	5	91.3

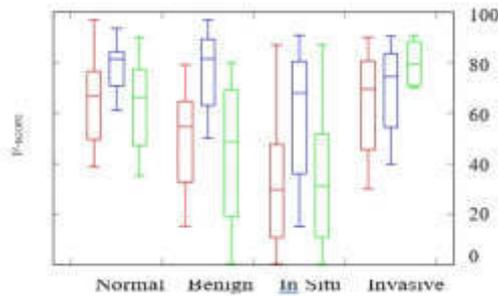


Fig. 4.2: F-Score comparative analysis

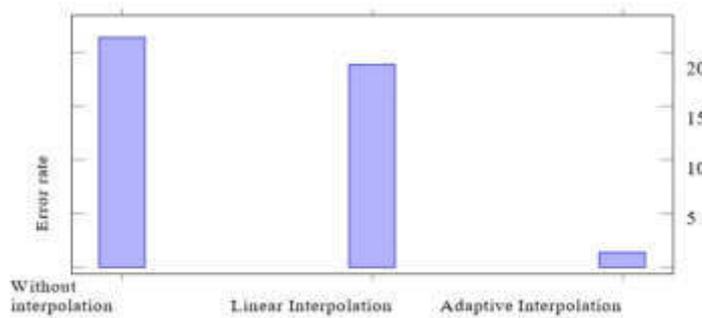


Fig. 4.3: Error rate analysis and without interpolation

Table 4.2 illustrates comparative analysis of proposed method with state-of-art approaches. Five different classes are considered to assess the accuracy of the system and the proposed system has achieved 91.3% accuracy than other approaches. Fuzzy-SVM model has also achieved notable accuracy which is around 90%.

Fig. 4.2, illustrates F1-score with box-plots of the proposed and state-of-art approaches over test set. All approaches have accomplished the most elevated F1 esteems due to effective clustering based on weight factor. The DL-AI approach has achieved essentially accurate outcomes for all test cases. Most researchers have used a modified interpolation technique for weight assessment and the KNN algorithm for data clustering to identify missing characteristics in a breast cancer dataset. Fig. 4.3 illustrates performance analysis of a proposed system with and without interpolation; the model without interpolation has achieved an average 21% error rate compared to linear interpolation. However, adaptive interpolation has achieved leq5% error rate on average. The cluster has omitted if any of the parameters are mismatched and when the confidence weight factor value is insufficient.

Initial simulation is carried out with normal and abnormal tissue data-set which can be observe in Fig. 4.4. Compare to GRNN, DL-AI has achieved high performance rate than FSVM in terms of accuracy 97% and sensitivity 98.3%, and specificity is 94% respectively. FSVM has occupied second position in performance rate but the specificity is 91% which is greater than GRNN. The GRNN has achieved 95% accuracy, 98.3%

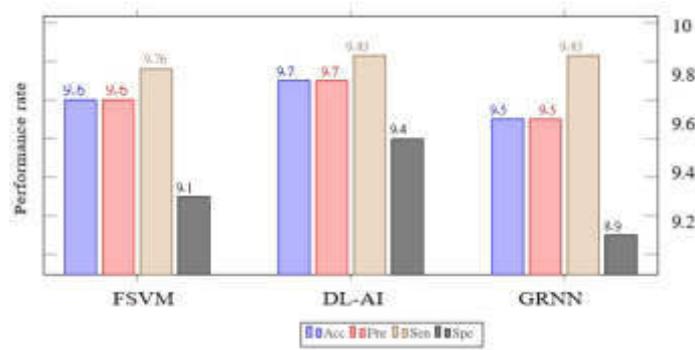


Fig. 4.4: Simulation-1 with normal and abnormal tissues

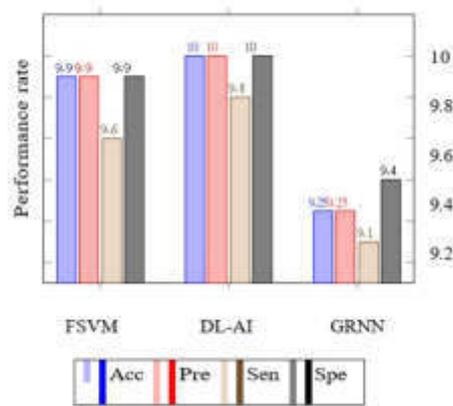


Fig. 4.5: Simulation-2 with benign and malignant

sensitivity, 89% specificity respectively.

Subsequently, Fig. 4.5 illustrates the simulation-2 results with benign and malignant tissue data. DL-AI approach has achieved greater performance than state-of-the-art approaches, like 100% accuracy, 98% sensitivity, and 100% specificity. However, GRNN has achieved low performance rate with the same tissue data set, where 92% accuracy, 91% sensitivity, 94% specificity are measured respectively. As usual, FSVM has held the second position which is better than GRNN.

Consequently, Fig. 4.6 illustrates simulation-3 result analysis report with invasive and in situ carcinoma tissue data. Here, FSVM has achieved low performance rate than GRNN such as 91% accuracy, 90% sensitivity, 92% specificity, but GRNN has achieved 92.5%, 91%, 93% rate towards accuracy, sensitivity, and specificity respectively. DL-AI has achieved its notable accuracy by 92%, 90% sensitivity, and 94.7% specificity respectively.

Table 4.3 illustrates standard deviation (SD), Area Under the Curve (AUC) measurement values to assess the performance of the system. In experiment-1, 500 normal and abnormal images are considered to estimate the SD, AUC based on three listed classifiers. Where, FSVM, GRNN, DL-AI have achieved 0.07%, 0.05%, 0.11% SD respectively. In experiment-2, FSVM, GRNN, DL-AI classifiers have achieved 0.19%, 0.11%, 0.21% of SD, and 0.97%, 0.99%, 1% of AUC respectively. Subsequently, this second simulation is carried out with 400 benign and malignant tissue data set. In experiment-3, FSVM, GRNN, DL-AI classifiers have achieved 0.15%, 0.163%, 0.20% of SD, and 0.98%, 0.967%, 0.99% of AUC respectively. In all three cases, DL-AI approach has achieved high performance rate than other approaches with different data-sets.

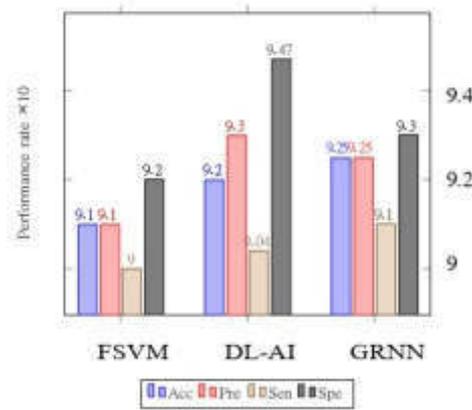


Fig. 4.6: Simulation-2 with invasive and in situ carcinoma

Table 4.3: Simulations results analysis with AUC and SD based on each classifier.

Experiment	Number of images	Classifier	AUC	SD
1	500	FSM	0.99	0.07
		GRNN	1	0.05
		DL-AI	0.99	0.11
2	400	FSM	0.97	0.19
		GRNN	0.99	0.11
		DL-AI	1	0.21
3	300	FSM	0.98	0.15
		GRNN	0.967	0.163
		DL-AI	0.99	0.2

**5. Conclusion.** The proposed system effectively determines breast cancer-based histopathology images to mitigate mistaken diagnosis. Deep learning networks accurately assessed the patterns for annotating histological data with a label system that enhances the system accuracy by 91.3%. The feature acquisition and missing analysis strategies have played an essential role in strategically achieving the targeted accuracy based on the entropy confidence weight factor. The attention-based rule combination effectively formulated the classification issue by detecting and predicting the different patterns to construct the classification rule. This process has optimized the clustering data by assessing missing parameters based on mean square error and data interpolation.

Note: KNN algorithm has been used during data clustering considering the above essential factors.

Simulation results show that the proposed system has achieved 91.3% accuracy than state-of-art approaches.

REFERENCES

- [1] HEATHER D COUTURE, LINDSAY A WILLIAMS, JOSEPH GERADTS, SARAH J NYANTE, EBONEE N BUTLER, JS MARRON, CHARLES M PEROU, MELISSA A TROESTER, AND MARC NIETHAMMER, *Image analysis with deep learning to predict breast cancer grade, ER status, histologic subtype, and intrinsic subtype*, in NPJ breast cancer, 4(1), 2018, pp. 30.
- [2] MAI BUI HUYNH THUY AND VINH TRUONG HOANG, *Fusing of deep learning, transfer learning and gan for breast cancer histopathological image classification*, in International Conference on Computer Science, Applied Mathematics and Applications, 2019, pp.255–266.
- [3] ZIXIAO LU AND ZHAN, *Brcaseq: a deep learning approach for tissue quantification and genomic correlations of histopathological images*, in Genomics, proteomics & bioinformatics, 2021.
- [4] PIN WANG, PUFEI LI, YONGMING LI, JIAXIN WANG, AND JIN XU, *Histopathological image classification based on cross-domain deep transferred feature fusion*, in Biomedical Signal Processing and Control, 68:102705, 2021.
- [5] ANDREAS HOLZINGER, BERND MALLE, PETER KIESEBERG, PETER M ROTH, HEIMO MÜLLER, ROBERTREIHS, AND KURT

- ZATLOUKAL, *Towards the augmented pathologist: Challenges of explainable ai in digital pathology*, arXiv preprint arXiv:1712.06657, 2017.
- [6] UDDARAJU, SUSMITHA, GP SARADHI VARMA, AND M. R. NARASINGARAO, *Prediction of NAC Response in Breast Cancer Patients Using Neural Network*, in Scalable Computing: Practice and Experience 23.4, 2022, pp: 211–224.
- [7] SCOTTY KWOK, *Multiclass classification of breast cancer in whole-slide images*, in International conference image analysis and recognition, Springer, 2018, pp: 931–940.
- [8] FRANCISCO MANUEL GASCA-SANCHEZ, SANDRA KARINA SANTUARIO-FACIO, ROCÍO ORTIZ-LÓPEZ, AUGUSTO ROJAS-MARTINEZ, GERARDO MANUEL MEJÍA-VELÁZQUEZ, ERICK MEINARDO GARZA-PEREZ, JOSÉ ASCENCIÓN HERNÁNDEZ-HERNÁNDEZ, ROSA DEL CARMEN LÓPEZ-SÁNCHEZ, SERVANDO CARDONA HUERTA, AND JESÚS SANTOS-GUZMAN, *Spatial interaction between breast cancer and environmental pollution in the monterrey metropolitan area*, in Heliyon, 7(9):e07915, 2021.
- [9] LORENA CONSOLINO, PIETRO IRRERA, FERIEL ROMDHANE, ANNASOFIA ANEMONE, AND DARIO LIVIO LONGO, *Investigating plasma volume expanders as novel macromolecular mri-contrast agents for tumor contrast-enhanced imaging*, in Magnetic Resonance in Medicine, 86(2), 2021, pp: 995–1007.
- [10] CHAOSHENG ZHANG, RENGUANG ZUO, YIHUI XIONG, XUN SHI, AND CONAN DONNELLY, *Gis, geostatistics, and machine learning in medical geology*, in In Practical Applications of Medical Geology, Springer, 2021, pp: 215–234.
- [11] ABBAS TAVASSOLI, YADOLLAH WAGHEI, AND ALIREZA NAZEMI, *Comparison of kriging and artificial neural network models for the prediction of spatial data*, in Journal of Statistical Computation and Simulation, 2021, pp: 1–18.
- [12] UDDARAJU, SUSMITHA, AND M. R. NARASINGARAO, *Predicting the Ductal Carcinoma Using Machine Learning Techniques—A Comparison*, in Journal of Computational and Theoretical Nanoscience, 16.5-6, 2019, pp: 1902–1907.
- [13] XIAOLING LENG, REXIDA JAPAER, HAIJIAN ZHANG, MILA YEERLAN, FUCHENG MA, AND JIANBING DING, *Relationship of shear wave elastography anisotropy with tumor stem cells and epithelialmesenchymal transition in breast cancer*, in BMC Medical Imaging, 21(1), 2021, pp: 1–16.
- [14] K VENKATA RATNA PRABHA, D VAISHALI, PALLIKONDA SARAH SUHASINI, K SUBHASHINI, AND P RAMESH, *Different diagnostic aids and the improved scope of establishing early breast cancer diagnosis*, in Micro-Electronics and Telecommunication Engineering, Springer, 2021, pp: 65–72.
- [15] LAILA AL-QAISI, MOHAMMAD A HASSONAH, MAHMOUD M AL-ZOUBI, AND ALA’M AL-ZOUBI, *A review of evolutionary data clustering algorithms for image segmentation*, in Evolutionary Data Clustering: Algorithms and Applications, 2021, pp: 201–214.
- [16] MARAM ALWOHAIBI, MALEK ALZAQEBAH, NOURA M ALOTAIBI, ABEER M ALZAHIRANI, AND MARIEM ZOUCHE, *A hybrid multi-stage learning technique based on brain storming optimization algorithm for breast cancer recurrence prediction*, in Journal of King Saud University-Computer and Information Sciences, 2021.
- [17] THOMAS HELLAND, BJØRN NAUME, STEINAR HUSTAD, ERSILIA BIFULCO, JAN TERJE KVALØY, ANNA BAR BRO SÆTTERSDAL, MARIT SYNNESTVEDT, TONE HOEL LENDE, BJØRNAR GILJE, INGVIL MJAALAND, ET AL. , *Low z-4ohtam concentrations are associated with adverse clinical outcome among early stage premenopausal breast cancer patients treated with adjuvant tamoxifen*, in Molecular oncology, 15(4), 2021, pp: 957–967.
- [18] FABIO A SPANHOL, LUIZ S OLIVEIRA, PAULO R CAVALIN, CAROLINE PETITJEAN, AND LAURENT HEUTTE, *Deep features for breast cancer histopathological image classification*, in 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC), IEEE, 2017, pp: 1868–1873.
- [19] UDDARAJU, SUSMITHA, AND M. NARASINGARAO, *A survey of machine learning techniques applied for breast cancer prediction*, in International Journal of Pure and Applied Mathematics 117.19, 2017, pp: 499–507.

*Edited by:* Polinpapilinho Katina

*Special issue on:* Scalable Dew Computing for Future Generation IoT systems

*Received:* Aug 18, 2023

*Accepted:* Dec 2, 2023



## SECURESENSE: ENHANCING PERSON VERIFICATION THROUGH MULTIMODAL BIOMETRICS FOR ROBUST AUTHENTICATION

SAMATHA J\* AND G .MADHAVI†

**Abstract.** Biometrics provide enhanced security and convenience compared to conventional methods of individual authentication. A more robust and effective method of individual authentication has emerged due to recent advancements in multimodal biometrics. Unimodal systems offer lower security and lack the robustness found in multimodal biometric systems. The research paper introduces a novel approach, employing multiple biometric modalities, including face, fingerprint, and iris, to authenticate users in a multimodal biometric system. The paper proposes the "Secure Sense" framework, which combines multiple biometric modalities to improve person verification accuracy. The proposed system utilizes both web-based and real-time datasets. For the web-based dataset, we employed the Chicago Face dataset for facial data, the MMU1 dataset for iris data, and the SOCOFing dataset for fingerprint data. In real-time data collection, facial data is captured using a Zebtronics Zeb-Gem webcam, fingerprint data is obtained using the Mantra MFS scanner, and iris data is collected using the Mantra MIS scanner. In the envisioned system, we introduce an innovative approach that employs a decision-level fusion technique across three unique biometric modalities, resulting in an impressive accuracy rate of approximately 93% across all modalities.

**Key words:** Biometrics, CNN, Daugman, Decision level fusion, Haar-Cascade, multimodal

**1. Introduction.** The process of identifying individuals based on their physiological, biological, or behavioral characteristics is known as biometrics[1]. These qualities are unique to each person and do not change more or less throughout life. Biometric security has proven to be an effective tool compared to electronic security. Biometric traits can be either physiological or behavioral traits of humans. This is possible as long as you have the following properties: Integrity, Uniqueness, Lasting Quality, Collectability, Circumvention, Value, Execution. Body type is a factor in physiological biometrics. Fingerprinting is a feature that has been around for over a century. Examples include irises, DNA, palm prints, and hand shapes. Behavioral biometrics relate to human behavior. The main trademarks used are brands that are still widely used today. Keystrokes, gait (the way you walk), and writing style are some additional factors. The ability to speak is one biometric feature that falls into both categories. Choosing the right biometric based on your application is important.

Sensor data is noisy and not universal, and in monomodal biometric systems using a single biometric, there is no fixed representation of the biometric. Discourse, for example, biometric element whose characteristics change radically when a person is affected by a cold or home conditions. Creating a multimodal biometric system that integrates evidence from multiple biometric sources can alleviate some of these problems. Multimodal or multi-biometric frameworks use a variety of physiological or social biometric data to record and identify people. An implementation of such a multimodal biometric personal verification framework is presented in this paper.

Some features offer great reliability and accuracy, but no 100% accurate biometrics. With the growing global need for security, there is a clear need for robust automatic person recognition systems. Accuracy of authentication is always of paramount importance when dealing with applications containing sensitive data. thus, facilitating multimodal biometrics. An integrated prototype system that includes different biometric types is called multi-biometric. Combining multimodal biometrics with personality verification can improve the performance of character verification frameworks. Multimodal biometrics can improve false acceptance rate (FAR) performance without reducing the likelihood of access denial by increasing the ability to distinguish between genuine and fraudulent classes[2]. Although there are many challenges in enrolling large populations using a single (unimodal) biometric, identifying unimodal biometrics is highly effective. The main problem with unimodal biometric systems is the fact that no single technology is suitable for all applications. As a result,

---

\*Department of CSE, Matrusri Engineering College, JNTUK, India ([samathajuluri@gmail.com](mailto:samathajuluri@gmail.com))

†Department of CSE, University College of Engineering, JNTUK, India ([madhavi.researchinfo@gmail.com](mailto:madhavi.researchinfo@gmail.com))

the use of multimodal biometric systems overcomes the drawbacks of unimodal biometric systems. Multi-biometrics are widely used all over the world. Various systems require reliable person recognition systems to verify or determine the identity of individuals. These schemes ensure that only authorized users can access the services offered [3]. Examples of such applications include secure access to buildings, computer systems, laptops, mobile phones, and ATMs [4]. These systems are vulnerable to fraudulent schemes in the absence of robust personal recognition engines. Due to unimodal limitations, an authentication scheme based on only one modality may not meet all requirements. Current work is expected to create a bimodal biometric framework using facial, fingerprint, and iris highlights. This is to mitigate the impact of certain obstacles in the unimodal biometric framework.

Robust authentication through multimodal biometrics entails the utilization of multiple biometric factors in conjunction to bolster security during user authentication. This approach offers a range of advantages, including heightened precision, enhanced security, increased resistance to spoofing attacks, and adaptability to varying environmental conditions.

**2. Related Works.** The multimodal biometric authentication system has been the subject of several proposals and developments. None of the biometrics is 100% accurate, despite the fact that some of the characteristics offer good performance in terms of reliability and accuracy. The authentication accuracy of the system is always the primary concern for applications that involve the flow of confidential information. Because of this fundamental reason, multimodal biometrics should be used [5].

Elhoseny [5] proposed the majority of security frameworks fall into one of these three categories: based on information; What you know, like a PIN, password, or identification, could very well be guessed, ignored, or shared. A token known as "What you have," such as a key or card, is another type; It could be taken, duplicated, or misplaced. The final type is the use of biometrics; what you are, like your face, IRIS, fingerprint, and other characteristics. Biometric identification systems can identify individuals by comparing a template set that is stored in the database to the measured and analyzed physiological or behavioral characteristics. Uni-modular biometric frameworks suffer from certain issues like noise in detected data, incompleteness, parody attacks, intra-class varieties, and similarity between classes. The utilization of a blend of at least two biometric types to further develop a framework's security is known as a multi-modular biometric framework. Models include Iris and fingerprints, to enhance user authentication and security. There are five levels of fusion in multi-modal biometric systems: scale sensor; where, at the feature level, the sensor's raw data are combined; At this level, the features of each user biometric process are combined to form a single feature set or score level; where the final decision, rank level, and match scores from difference matches are combined to show how similar the input and stored templates are to one another; Each selected personality is assigned a position by each biometric subsystem, and the subsystem positions are combined to create a new position for each character and choice level; The last acknowledgment choice is made by consolidating the aftereffects of each biometric subsystem. Multi-biometric systems can be broken down into six distinct groups: diverse sensors; which use a variety of sensors and algorithms to capture biometric characteristics and extract various data; where numerous cases of the equivalent biometric information are handled by various calculations; use numerous examples, different occurrences of the equivalent biometric, for example, left and right irises or pointers; In order to obtain a more comprehensive multi-modal representation of the underlying biometric, multiple samples of the same biometric are taken using the same sensor; hybrid, in which evidence from two or more biometric features is combined; refers to systems that use two or more of the five other categories. This chapter presents a proposed system for fingerprint and iris recognition that is based on minutiae extraction for fingerprint recognition and hamming distance for IRIS Recognition. The proposed system is built with MATLAB 7.8.0.347(R2009a) and makes use of CASIA Iris V1 data for Iris recognition and FVC 2000 and 2002 DB1 A data for fingerprint recognition. It compares the standalone fingerprint recognition system's FAR, FRR, and accuracy metrics with those of the multi-modal biometric system based on Fingerprint and Iris, demonstrating that the multi-modal system's FAR and FRR results are lower and accuracy is higher than those of the standalone fingerprint recognition system [6]. The experiment's findings were based on Iris recognition datasets from CASIA Iris V1 and fingerprint recognition datasets from FVC 2000 and 2002 DB1 A.

The author in [7] explained, Feature Extraction, and Classification are the three distinct phases of this novel biometric image classification approach, which is based on an optimal neural network and presented in

this paper. The novel method begins with the procedure of pre-processing the median filter. The element extraction stage, in which different highlights are really separated from biometric pictures, comes straightaway. Shape and surface components like the finger impression and ear include are among the removed highlights. The ONN method does a great job of classifying the images. The sensitivity, specificity, and accuracy of the new method's efficiency metrics, as well as the False Positive Rate, False Negative Rate, and accuracy, are successfully estimated. It breezes through without a hitch with regards to grouping pictures with excellent viability, accomplishing real effectiveness in the errand of characterizing pictures, and delivering energizing outcomes with extraordinary exactness.

The development of sensing and multi-modal communication technologies has resulted in an increase in the complexity of multi-modal biometric recognition (MBR). The presence of biometric data with higher dimensionality and more diverse inputs has also added complexity to MBR. This paper proposes a multi-modal fusion technique system based on the GLCM algorithm's combination of features from fingerprint and iris images. Prior to suggesting the proposed acknowledgment framework, a survey of past investigations on the iris and unique mark was finished. The AND entryway was used to pursue the final choice in the combination method choice. The review's results showed that, in comparison to the KNN classifier's recommended limit of up to 90%, the proposed framework achieved a high accuracy rate. The system's evaluation was based on the FAR, FRR, and total accuracy rate [8].

Application security cannot be guaranteed without authentication. Cloud computing is a model for providing IT-related services to a variety of end users that uses the internet. Cloud services are becoming easier to use over time because they give users a lot of freedom. Also, it raises a few worries in regard to information security. The inherent biological patterns of the multi-modal biometric system enhance the robustness of the authentication mechanism. Because of the caught design, it precisely separates the people from their characteristics. Additionally, this concept can be applied in a variety of ways to strengthen the system, including, among other things, securing health information and the human genetic code for future reference through digital ledger management and EHR management. This work proposes a multi-modal biometric authentication method to increase cloud-based data security. An MD-5 hashing calculation is used to combine the elements removed from a unique finger, iris, and palm print in multiple stages to produce a hash of strings and numbers for a mystery key. The protected data is encrypted with one of three symmetric key encryption algorithms—DES, AES, or Blowfish—using the secret key. AES outperforms the other two algorithms in terms of performance in relation to the strength of the encryption process, while DES takes less time to execute. This model demonstrated its robustness in data security [9] thanks to the fusion of human modalities in the security mechanism.

Paper [5] proposes CNN-based deep learning models for the feature-level fusion of online fingerprints and signatures. At convolution and fully connected layers, early and late feature fusion techniques that combine features from both biometric modalities have been developed. The fingerprint input image has a fixed size of  $150 \times 150 \times 1$ , and the online signature file is  $17 \times 17$  pixels. The size of the signature was changed to  $1 \times 17 \times 1$  before it was sent to the network for online signatures in order to combine the characteristics of an online signature and a fingerprint. Different systems were attempted to consolidate the elements of an internet-based signature and a picture of a unique finger impression. Be that as it may, the width of the web-based mark's component vector, which was set to 1, didn't work on the precision or some other qualities for other assessment measurements for the proposed framework. The problem was solved, and the system's accuracy and values for other evaluation metrics were raised, by adding two zero-padded layers to the signature network. The additional zeros were added, at least at the top, base, left, and right, of the element vector using this zero-cushioning strategy. As a result, the dimensions of the final feature vector increased to  $4 \times 4$ . Similarly, the final feature vector for fingerprints was  $4 \times 4$ . These features have been concatenated and passed through fully connected layers in order to extract and classify features in a more abstract way. The model was trained and tested using the new data set. With the early fusion scheme, the system was 99.10% percent accurate, and with the late fusion scheme, it was 98.35% accurate. The fusion may also use low-level fingerprint characteristics or level 3 features, such as active sweat pores and ridge contours, in the future to ensure a user's accuracy and liveliness. One of the various biometrics-specific different state-of-the-art cryptography methods may be incorporated into the proposed system in the future to further ensure the fused biometric template's safety.

Paper [6] proposes a useful feature-level fusion method for a multimodal biometric recognition system. We

looked at the multimodal biometric, which includes a level combination like an ear and palm impression. The four main steps in our proposed method are preprocessing, feature extraction, optimal feature level fusion, and recognition. The shape highlights were separated utilizing a changed district developing calculation, and the surface elements were extricated utilizing the HMSB administrator. Additionally, we selected the relevant features by employing the optimization technique. For picking the ideal component, we used the OGWO + LQ computation. On the positive, we proposed affirmation, for the affirmation we used the multi-part support vector machine (MKSVM) estimation. Measures like responsiveness, particularity, and precision are used to evaluate the implementation of our proposed strategy. Our proposed strategy outperforms other techniques in terms of awareness, particularity, and precision, as demonstrated by the trial results and similar research. Consequently, the multimodal biometric recognition system can greatly benefit from the efficiency of our suggested method. Individual confirmation demonstrates utilizing convolution brain networks is an effective method for putting a multimodal biometric framework on equipment. Our proposed work's execution on CNN [22] would be a future undertaking.

**3. Existing Unimodal based Systems.** Unimodal systems are biometric identification systems that use a single biometric trait of the individual for identification and verification[10].

*Fingerprint Recognition.* During this recognition process, fingerprint features like swirls, arcs, loops, raised patterns, ridges, and other fine details are recorded using ink or a digital scanner to take an image of the fingerprint[11]. Then, we store or process this data with cryptographic algorithms. Particulars can be planned to relative finger positions utilizing a product program, and comparable details data can be looked through in a data set. To speed up the search, images are converted into strings. So, in most cases, the fingerprint image is just a string of characters to compare. By holding their finger on the reader for a few seconds, fingerprint readers enable users to identify or verify an individual. Today, fingerprint readers also check for blood flow and make sure the ridges on the ends of the fingers are where they should be to make sure no fake fingers are used. One of the benefits of this procedure is that it makes use of advanced technology and is easy to use. In addition to supporting the capability of enrolling multiple fingers for enhanced anti-spoofing properties, this method is extremely accurate and stable over time[9]. There are some drawbacks to this method: Contacting the unique finger impression per user can smirch your fingers, which can influence picture quality. The skin type and condition may influence the registered data [28].

*Facial Recognition.* This identification interaction estimates the whole facial design, like the distance between the eyes, nose, mouth, and the edge of the jaw [12]. Functions are used to create and store templates in the database. A single image is formatted based on facial features and captured with a conventional or video camera during a confirmation or differential verification cycle. This format is then contrasted and the layout of the data set. In the future, only high-performance test procedures will make use of this acknowledgment arrangement. Before authentication, systems today demand that users smile, blink, and move in a humane manner[34]. This is done to stop people from pretending to be part of the system. This framework enjoys the benefit of being inconspicuous. It functions admirably at an insignificant cost as it permits you to remove pictures utilizing normal reconnaissance cameras. Additionally, this strategy can be used covertly in high-security settings like airports. To put it another way, the person is not aware that the picture was taken. The disadvantage is that appearance and the environment can have an impact on facial recognition, which is highly dependent on the quality of the images obtained. This system can cause issues with identical twins and is typically inaccurate. Additionally, it operates secretly, jeopardizing your privacy[13].

*Hand Geometry.* Utilizing this recognition system is a straightforward but highly accurate process. The user places their hands on a metal surface with guidance pegs to assist in proper hand alignment. The gadget finds out around 90 hand qualities, including length, width, thickness, finger surface region, and palm size. These characteristics can be checked against another template in the database or form a template that is stored in the database. The fact that hand geometry technology has one of the biometrics industry's smallest templates—typically less than 10 bytes—is one of its primary advantages [12]. It likewise has high client mindfulness and is non-nosy. The searcher's relatively large size and low precision are its weaknesses. Children, people whose joints are inflamed, people whose fingers are missing, or people with large hands may find reporting challenging. Now, it is only used for the verification process.

*Iris recognition.* To identify individuals, this technology looks for characteristics like freckles, furrows, and rings in the colored tissue around the pupil [24]. This is one of the biometric technologies with the highest degree of accuracy. As indicated by a new Silicon.com article, an administration exploration including facial acknowledgment, iris acknowledgment, and unique finger impression acknowledgment found that iris acknowledgment played out the best as a confirmation technique, in spite of the difficult enrolment process. A normal camcorder can be utilized to get a picture of the iris, and it very well may be finished from a more noteworthy distance than with a retinal output. Users of this technology need to work together to get a clear picture. Subsequently, the gadget is planned so that when the client puts his head before it, he can see his iris reflecting in the gadget, demonstrating that an unmistakable picture can be gotten. In order to ensure that the framework is not fooled by a fictitious eye, the device may alter the light that is reflected into the eye and observe the pupil expanding. This framework has the advantage of a high confirmation rate and fraud protection. The iris data are the same for the left and right eyes and do not change as you get older. The disadvantage of this strategy is that it is extremely intrusive; The process of enrolling is somewhat challenging since not everyone is familiar with the system. John Daugman, a Cambridge College scientist, was responsible for a great deal of exploration on iris acknowledgment.

*Retinal Recognition.* It is a technique for observing the retina's blood vessel pattern, which is said to vary from person to person. Essentially, the retina is comprised of material tissue with various layers close by photoreceptors like cones and shafts. Photoreceptors take in the light rays they emit and turn them into an electrical force that the mind uses to create images. In comparison to the tissue that surrounds the eye, the blood vessels that make up the retina are better able to detect and reflect infrared light. In this way, the gadget utilizes its IR light to enlighten the retina, and when the IR is mirrored, a retina assessment gadget utilizes it to eliminate the remarkable components of the retina utilizing different estimations. The size of the pupil, which controls how much light reaches the retina, affects image quality. The fact that a person's retina stays the same throughout their lifetime is a benefit of this innovation. In comparison to other technologies, it has many distinctive features and a high verification rate. Client anxiety when eyes are filtered from very close distances, the requirement to remove eyeglasses at contact hotspots, and the widespread perception that devices can damage the retina are all obstacles to this strategy. Being patient and cooperative with a lot of customers [27-29].

*Signature Recognition.* The behavioral part of how we sign names is analyzed by Mark's Acknowledgment. This development is due to direct qualities such as timing, pressure, speed, overall changes, and different stroke paths throughout the stamping. Despite all the characteristics of copy brands being fundamental from the ground up, the nature of their behavior is difficult to copy. This device includes a stylus (or pointer) and a There are special writing tablets that The pen must be used by the customer to sign on to the tablet. This system collects all information about features, creates templates, and stores them in a database. You are required to register for various events to complete valid registration. This method is advantageous because it is a painless device that is well-known and difficult to imitate. The drawback is that the signature should not be too long or too short. A signature that is too long can contain too much behavioral data and make it difficult for the system to find consistent and unique data points. If the signature is too short, there may not be enough data points for the system to create a unique template. Whether you're standing, sitting, or resting your arms, the recording process should take place in the same environment. The structure also tries to modify the client, resulting in signature anomalies [27].

*Voice Recognition.* Voice originates from the vocal cords. When we try to communicate, the distance between the vocal cords narrows or widens. Thus, increasing expansion with deep breathing and narrowing with exhalation produces an unmistakable sound. Most of the vocal tract consists of the larynx, oropharynx, oral cavity, nasopharynx, and nasal cavity. In this method, users are asked to recite part of a list of texts or numbers using a microphone or phone connected to their computer. A computer receives the user's voice and converts it into digital form. A template is created by saving this format and extracting certain features from it. Statistical profiles are created by comparing multiple samples and identifying various recurrence patterns. As a result, statistical profiles are compared to ensure accuracy. The advantage of this strategy is that existing communication bases or mouthpieces can be used. It is unobtrusive and easy to use. This is a problem because background noise affects the system. Accuracy is very low, as it can be affected by voice adaptation due to

aging, disease, alcohol consumption etc [23].

*Gait Recognition.* Gait recognition is a technology that recognizes people by the unique way they walk. Stride recognition is usually useful for reconnaissance because it tends to be caught without the person's consent. Faking and hiding are also extremely challenging. Energizers like liquor and medications, which make individuals unequal, actual changes brought about by pregnancy, a mishap, weight gain or misfortune, an individual's state of mind, and the garments they wear can all influence walking. Stride acknowledgment should be possible by utilizing typical observation cameras. A portion of the benefits are that it tends to be finished in a good way and minimal expense observation cameras can be utilized. It is difficult to fake and can be affected by the person's background, clothes, and emotions, which are disadvantages [27].

*Key stroke Recognition.* Because different people type differently on computer keyboards, keystroke detection is necessary. This strategy includes a keyboard or a computer-attached keyboard. When asked to enter a particular word on multiple occasions, the user must meet certain requirements, such as typing correctly. In the event of typos, You must start from scratch. The manner in which an individual sorts, the total composing speed, the time between progressive keystrokes, how long each key is held down, and how frequently an individual presses other keys utilizes the console. The request wherein capitalized letters are placed, for example, the numeric keypad and capability keys, whether the shift key or the letter key is delivered first, and so on., are a few of the extracted characteristics. These characteristics are used to generate a template for a statistical profile of a person's behavioral characteristics. They are advantageous because they are entirely software-based, do not require additional hardware, integrate with other biometrics, and require little training. Another advantage is that you can change the word and make a new template if you change the template for a particular word. This method has one drawback: it doesn't make it easier to remember passwords. In addition, the technology is still in its infancy and has not yet undergone a significant amount of testing.

*Hand Vascular Pattern Recognition.* The technology uses state-of-the-art recognition algorithms based on the unique veins and capillaries found on the back of the human hand to validate or recognize human users. Vein pattern recognition is based on the concept that the hotter the image, the brighter it becomes during infrared imaging. The colder the object, the darker it appears. So, when you scan your hand with an infrared reader, your vein pattern will appear darker than the surroundings. Everyone's pattern is different. If the image is saved as a personal template, it can be used for verification or identification purposes. The advantage is the high accuracy of this method. Unaffected by harsh environments such as construction sites, military bases, and manufacturing plants [13]. It is also very convenient for users as it requires minimal knowledge of the system. Infrared absorption patterns can be easily compared via optical and DSP techniques to provide a large, robust, and stable base for matching units. This usually requires a lower-resolution IR [27].

### 3.1. Limitations of Unimodal Systems.

The following are the limitations [24] of unimodal systems.

*Distortion of the biometric data that was input.* Users may be rejected or identified incorrectly as a result of distorted biometric data preventing the alignment process with database templates.

*Intra class variations.* The biometric information obtained during identity verification may differ from the information used to create the registration format, which may affect verification systems. Biometric forms should show slight differences within the class.

*Interclass similarities.* Biometric highlights ought to be altogether unique for various individuals and ought to guarantee little similitudes between classes in the component space. Any biometric system can only effectively distinguish a certain number of users. The limitations of the recognizable proof framework cannot be arbitrarily extended to calculus consistent with a fixed set of element vectors.

*Non-universality.* There should be a lot of variation between classes in the biometric template. It is not always possible to obtain accurate and useful biometric data from users.

*Intruder attacks.* This type of attack modifies biometric authentication functionality to avoid detection. You could even create a fake biometric design to recognize someone else's personality.

**4. Proposed System.** In the proposed system, we incorporate three distinct biometric modalities: face, iris, and fingerprint. Our dataset comprises both web-based and real-time data sources. The web-based data includes well-established online datasets, specifically, the Chicago face dataset for face recognition, the MMU1 dataset for iris recognition, and the SOCOing dataset for fingerprint recognition. In contrast, the real-time dataset is gathered using webcam technology as well as Mantra MIS and MFS scanners. Subsequently, we

conduct feature extraction on all these modalities. We employ decision-level fusion techniques to amalgamate these extracted features. Finally, the authentication process is carried out using a Convolutional Neural Network (CNN) algorithm.

#### 4.1. Datasets.

##### 4.1.1. Web Based Dataset.

*Face Dataset.* The Chicago Face Database was created by Bernd Wittenbrink, Debbie S. Mom, and Joshua Correll at the School of Chicago. It offers standard, high-resolution portraits of men and women whose ethnicities range from 17 to 65. There are a lot of norming data for each model [14-16].

*IRIS Dataset.* The Multimedia University (MMU1) database contains eye images for the IRIS-based biometric system's training models. Each person's individual IRIS patterns for each eye make it easier to identify them. 46 people's left and right IRIS are represented by 460 images and a few empty files in this dataset. Individual ID and characterization of an IRIS picture in view of a saved information base can be achieved through IRIS division [17].

*Fingerprint Dataset.* A biometric fingerprint data set produced for scholarly exploration is the Sokoto Coventry Finger Impression Dataset (SOCOFing). SOCOFing is comprised of 6,000 fingerprint images from 600 African subjects. These fingerprint images include synthetically altered versions with three distinct levels of alteration—obliteration, central rotation, and z-cut—as well as distinctive features like labels for gender, hand and finger name, and so forth [18].

**4.1.2. Real Time Dataset.** We have gathered the real-time dataset for our task, from the students. We have taken ten images, four of each student's face, three images of their fingerprints, and three images of their iris.

The face pictures are gathered utilizing the Zebronics Zeb-Gem Genius webcam. A USB-powered web camera with a 3P lens and clear videos is the Zeb-Crystal Pro. It has night vision, a clip-on design for easy mounting, and a built-in microphone. 30 frames per second, 640 x 480 video resolution, and 1.2 meters of cable. The webcam can be used directly without the need for drivers.

The fingerprint pictures are collected utilizing the Mantra MFS scanner. The MFS100 USB fingerprint sensor is a high-quality option for desktop or network security fingerprint authentication. MFS100 depends on optical detecting innovation which effectively perceives low-quality fingerprints too. MFS100 can be utilized for validation, ID, and check works that let your unique finger impression carry on like computerized passwords that cannot be lost, neglected, or taken. Scratches, impacts, vibration, and electrostatic shock are all prevented by a hard optical sensor. Attachment and play USB 2.0 rapid point of interaction upholds numerous gadgets taking care of. 500 dpi optical finger impression sensor scratch-free sensor surface. supports Windows 7, 8, 10, Vista, Windows 2000, Linux, Windows ME, and Windows 98 SE SDK, Libraries, and Drivers on all platforms. 32-Bit and 64-Bit Support for applications and easy integration into production servers [19]. This scanner needs MFS100 RD service and MFS100 driver to be downloaded.

The Mantra MIS scanner is used to collect the iris images. The Mantra MIS100V2 IRIS Scanner is a powerful and durable iris scanner that has been used in numerous projects and is renowned for its precision. MIS100V2 is an excellent USB IRIS Sensor for IRIS Confirmation to get to the work area or organization security. The Single IRIS Sensor MIS100V2 can be extensively utilized for banking, access control, and identity applications like Aadhaar Authentication. The inbuilt LED indications make it simple to adjust the MIS100V2 quality algorithm, which is able to quickly and easily identify low-quality IRIS images. The MIS100V2's fast auto-capture capabilities are due to its proprietary distance sensing and focus analysis technology. Downloading the MIS100V2 RD service and driver is required for this scanner [20].

#### 4.2. Feature Extraction Techniques.

**4.2.1. HAAR-CASCADE API:** Haar Cascade is an algorithm for detecting features in images. A cascade function is trained on a lot of positive and negative images for detection. The algorithm doesn't need a lot of work and can run in real-time. We can train our own cascade function for custom objects like cars, bikes, and animals. Since it just distinguishes the matching shape and size, Haar Outpouring can't be utilized for face acknowledgment. The flowing window and outpouring capability are used in Haar overflow. It makes an effort

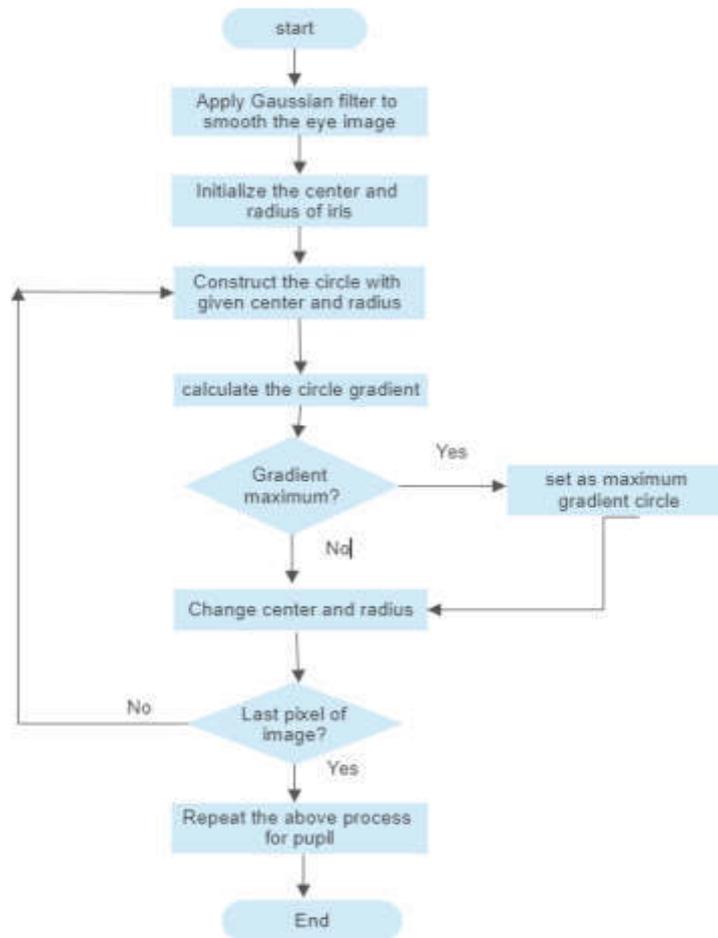


Fig. 4.1: Flow chart describing Daugman's method

to identify strengths and weaknesses for each window. Positive if the window is a component of something else; negative in this case. The Haar cascade can be used with a binary classifier. It gives the outpouring windows, which could be a piece of our article, a positive worth. additionally, detrimental to those windows that cannot be a component of our product. The abundance of pre-trained haar cascade files makes implementation extremely simple. We also have the ability to train our very own haar cascade, but for that to work, we need a lot of data. A GitHub repository managed by the OpenCV library contains all of the well-known pre-trained haars cascade files. These documents can be utilized for an assortment of item recognition undertakings, for example, Eye detection, vehicle detection, nose/mouth detection, body detection, and license plate detection are all methods of human face detection [23].

**4.2.2. DAUGMAN Algorithm for Iris.** The first step is to apply a Gaussian filter to the eye image. This helps in reducing noise and creating a smoother image. Initialize the center and radius of the iris. These parameters will be used as starting points for constructing the iris boundary. Using the initialized center and radius, you create a circle within the eye image. This circle represents the estimated boundary of the iris. Then calculate the gradient along the circle's boundary. This step helps identify edges and patterns in the iris. check if the gradient along the circle is at its maximum. If it is, this circle is considered as the maximum gradient circle, which corresponds to the iris boundary. If not, adjust the center and radius of the circle to refine the boundary estimation. After processing the entire circle, check if reached the last pixel of the image. If not go

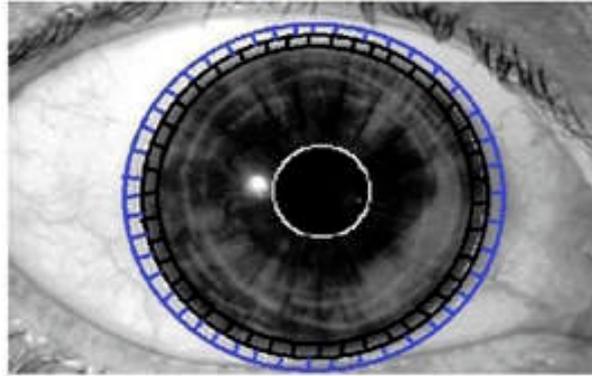


Fig. 4.2: The dark boxed line addresses a one-pixel wide iris line at a specific span 'R' (not to scale) at the middle direction and the blue boxed line addresses one-pixel wide circle at a similar focus coordinate at sweep 'R+1'

back to the step of constructing a circle with the adjusted center and radius, repeating the process for the next region of the iris. Repeat the Above Process for Pupil as shown in Fig 4.1.

**4.2.3. DAUGMAN'S Operator.** The task is to find the middle directions and the span of the iris and the student and Daugman's condition is utilized for this assignment. The integral of a differential equation is Daugman's theory of border recognition's central focus [23].

$$\max(r, x_0, y_0) \left| G_\sigma(r) * \frac{\partial}{\partial r} \oint_{r, x_0, y_0} \frac{l(x, y)}{2\pi r} ds \right| \quad (4.1)$$

The intensity of the pixel at coordinates  $(x, y)$  in the image of an iris is represented by  $l(x, y)$ .  $r$  signifies the sweep of different round areas with the middle directions at  $(x_0, y_0)$ .  $\sigma$  is the Gaussian distribution's standard deviation. A scale sigma Gaussian filter is denoted by the symbol  $G_\sigma(r)$ .  $(x_0, y_0)$  is the accepted focus direction of the iris. The parameters  $(r, x_0, y_0)$  specify the circle's contour as  $s$ .

The Daugman's operator in Equ. (4.1) is pivotal in improving iris features, ultimately resulting in precise user authentication.

By altering the center  $(x, y)$  of the circular contour and the radius  $(r)$ , the operator seeks the circular path with the greatest change in pixel values. In order to achieve precise localization, the operator is applied iteratively while the smoothing is gradually reduced. On the off chance that the factors  $x, y$  and  $r$  have a place with the reaches  $[0; X]$ ,  $[0; Y]$   $[0; R]$  separately, this strategy has the computational intricacy of the request  $[X * Y * R]$ . As a result, in order to compute the circle parameters using this method, a total of R scans are required at each pixel.

An inquiry over the whole picture (of an eye) is finished, pixel by pixel as depicted in Fig. 4.2. The normalized sum of all circumferential pixel values increases in radius at every pixel. At each degree of expanded span, the contrast between the standardized amounts of pixel power values at neighboring radii circles is noted. That pixel is deemed to be the center iris pixel when, following the entire search, summation, and differentiation portion of the calculation, the greatest variation in the sum of circumferential pixel intensity values exists between two adjacent contours.

**4.2.4. 2D Gaussian Filter.** In signal processing and electronics, a Gaussian filter is one whose impulse response is a Gaussian function. A Gaussian channel changes the info signal numerically through convolution with a Gaussian capability. According to our hypothesis, the Gaussian smoothing administrator is a 2-D convolution administrator used to smooth eye images to remove noise. The one-layered Gaussian channel has

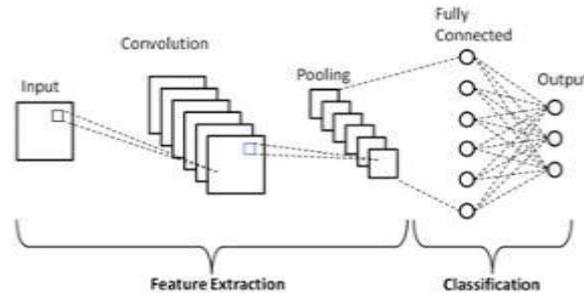


Fig. 4.3: CNN Architecture

a drive response given in Equ. (4.2)

$$g(x, y) = \frac{1}{2\pi\sigma^2} \cdot e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (4.2)$$

where  $x$  and  $y$  signify the separation from the beginning towards the edges of the channel in the level hub and vertical pivot bearings, separately and  $\sigma$  is the standard deviation of the Gaussian dispersion.

**4.2.5. 2D Convolution.** In math and useful examination, convolution is a numerical procedure on two capabilities  $f$  and  $g$  as given in Equ. (4.3), creating a third capability that is ordinarily seen as a changed variant of one of the first capabilities

$$(f * g)(t) = \int_{-\infty}^{\infty} f(r)g(t-r) dr \quad (4.3)$$

While the image  $t$  is utilized above, it need not address the time-space.

A 2D Gaussian filter is employed to achieve image smoothing or blurring, while a 2D convolution filter serves various image processing functions, including edge detection and feature extraction [33].

**4.3. Convolutional Neural Network.** CNNs are a subset of Profound Brain Organizations that are appropriate for visual picture examination since they can recognize and group picture highlights. PC vision, regular language handling, picture and video acknowledgment, and clinical picture investigation are only a couple of their many purposes. CNN is useful for image recognition due to its high accuracy. Medical image analysis, phone security, and recommendation systems are just a few of the many uses for image recognition [25]. The term "Convolution" is used in CNN to refer to the mathematical function of convolution. Convolution is a unique type of linear operation in which two functions are multiplied to create a third function that explains how one function's shape is altered by another [21]. In layman's terms, the output is created by multiplying two images that can be represented as matrices in order to extract features from an image [26]. A convolution device for isolating and recognizing the different picture highlights for use in highlight extraction examination. The feature extraction network contains numerous pairs of convolutional or pooling layers. a fully connected layer that, based on the features that were extracted in earlier stages, uses the convolution output to predict the class of the image. This CNN model of component extraction intends to diminish the number of features present in a dataset[21]. It adds new highlights to a current arrangement of elements, to sum up those highlights. There are a number of layers in the CNN architecture diagram.

There are three distinct types of layers in the CNN as depicted in Fig. 4.3 [29] pooling layers, fully connected (FC) layers, and convolutional layers. The stacking of these layers will result in the formation of a CNN architecture. The dropout layer and the activation function are two additional significant parameters that are described below in addition to these three layers [25].

**4.3.1. Convolutional layer.** The primary layer that is used to distinguish the various highlights from the information pictures is this one. The dot product (MxM) between the parts of the input image that are

proportional to the size of the filter is taken by sliding the filter over the image. The Feature map is the output, and it contains information about the image, like its corners and edges. The numerical activity of convolution is completed in this layer between the information picture and a channel of a specific size  $M \times M$ . Subsequently, this component map is dealt with in various layers to get to know one or two features of the data picture. The CNN convolution layer sends the result to the subsequent layer after applying the convolution operation to the input. Convolutional layers in CNN offer significant advantages because they ensure that the pixel's spatial relationship remains unchanged.

**4.3.2. Pooling Layer.** A Convolutional Layer is typically followed by a Pooling Layer. This layer's essential goal is to reduce computational expenses by making the convolved include map more modest. This is accomplished by reducing connections between layers and independently working on each component map. There are many different Pooling operations, each with its own set of rules and methods. It basically summarizes the features created by a convolution layer. In Max Pooling, the choice of the largest element is made using the feature map. The normal of the components in a picture part of a foreordained size is determined utilizing Normal Pooling. The total sum of the elements in the predefined section is computed using Sum Pooling. The Pooling Layer for the most part fills in as a framework between the Convolutional Layer and the FC Layer. This CNN model makes it easier for the networks to independently recognize the features because it generalizes the features that were extracted by the convolution layer. This also results in a reduction in computation time within a network.

**4.3.3. Fully Connected Layer.** The Fully Connected (FC) layer, which is used to connect the neurons between two distinct layers, is made up of the weights, biases, and neurons. These layers make up the last few layers in the majority of CNN architectures and come before the output layer. The flattened input image from the layers before it is received by the FC layer. After the vector has been smoothed, it goes through a couple of more FC layers, which are ordinarily where numerical capabilities tasks occur. At this point, the classification process begins. The fact that two completely associated layers will perform better than a single associated layer is the reason why two layers are associated. The amount of human oversight in CNN is reduced by these layers.

**4.3.4. Dropout Layer.** Overfitting in the preparation dataset is normal when all highlights are associated with the FC layer. When a model performs well on one set of data but fails to perform as well when applied to another set of data, this phenomenon is known as overfitting. To beat this issue, a dropout layer is utilized wherein two or three neurons are dropped from the cerebrum network during the planning process achieving a lessened size of the model. 30% of the hubs leave the brain organization haphazardly after passing a dropout of 0.3. An AI model's exhibition is improved by dropout since it works on the organization and forestalls overfitting. It removes neurons from the neural networks during training.

**4.3.5. Activation Functions.** Finally, the CNN model's enactment capability is one of its main limitations. They are used to estimate and learn about any kind of constant and intricate connection between the organization's factors. Simply put, it determines which model information should fire in the forward direction and which should not at the end of the network. The organization gains non-linearity accordingly. The ReLU, Softmax, tanH, and Sigmoid capabilities are only a couple of the enactment works that are much of the time used. Each of these capabilities serves a specific purpose. A CNN model typically employs softmax for multi-class classification, while the sigmoid and softmax functions are preferred for binary classification. In layman's terms, a CNN model's enactment capabilities decide if a neuron should be actuated. It comes to a conclusion about whether or not the contribution to the work is necessary to anticipate using numerical tasks.

Fig. 4.4 demonstrates the connections that exist between the various components that are being utilized. They are typically created with the intention of deepening one's comprehension of how the various parts interact with one another to accomplish the objectives.

#### 4.4. Workflow.

**4.4.1. Acquisition.** The primary module that collects biometric data from individuals is the hardware for biometrics sensor hardware. A camera sensor, for instance, captures images of the face and iris; a fingerprint scanner for fingerprint, and so on. The acquisition of iris and fingerprint images in our project is performed

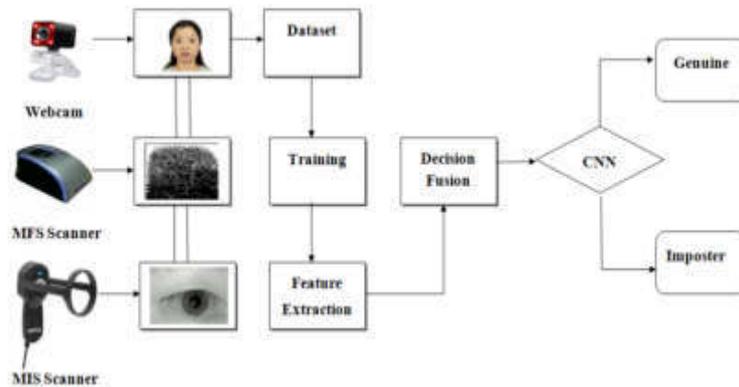


Fig. 4.4: Proposed System

using Mantra devices. We have used an MFS100 scanner for fingerprint acquisition, a MIS100V2 scanner for iris acquisition, and a ZERBRONICS webcam for the face.

**4.4.2. Dataset and Training.** The project makes use of both real-time and internet-based datasets. Chicago face dataset (CFD-INDIA)[16], MMU1 iris dataset[17], and Sokoto Coventry Fingerprint Dataset (SOCOFing) [18] make up the internet dataset. The mantra devices and webcam are used to collect the real-time dataset. These datasets are used to train the multimodal system.

**4.4.3. Feature extraction.** The captured biometric data is pre-processed to remove any additional imperfections or noises, and then it is processed further for the feature extraction procedure, which aims to select biometric features that are most likely to convey an individual's uniqueness.

#### 4.4.4. Preprocessing steps.

*Normalization.* The range of pixel intensity values is altered by this procedure. The aim of normalization is to bring the image into a range that is easy to sense. OpenCV involves standardized capability for picture standardization.

*Skew Correction.* When scanning or taking a picture of any document, there is a chance that the resulting image may occasionally be slightly skewed. The image's skewness should be corrected in order to improve the algorithm's performance.

*Image Scaling.* The image should have more than 300 PPI (pixels per inch) for better performance. Thus, assuming that the picture size is under 300 PPI, we really want to increment it. We can involve the Pillow Library in this.

*Noise Removal.* For smoothing the image, this step removes the small dots and patches that are more intense than the rest of the image. That is simple to accomplish with OpenCV's rapid N1 Means Denoising Colored function.

*Thinning and Skeletonization.* This step is performed for the manually written text, as various writers utilize different stroke widths to compose. The width of the strokes is uniformized with this step. OpenCV can be used for this.

*Gray Scale image.* An image moves from other color spaces to gray tones through this process. Between complete black and complete white, the color varies. OpenCV's `cvtColor()` capability plays out this undertaking without any problem.

*Thresholding or Binarization.* In this step, any colored image is transformed into a binary image with only black and white as colors. This can be accomplished by establishing a threshold, which is typically half of the 127-pixel range from 0 to 255. On the off chance that pixel esteem is more prominent than the edge, it turns into a white pixel; If not, it will turn into a black pixel. Otsu's Binarization and Versatile Binarization may be a superior choice for deciding the picture explicit limit esteem. The preprocessing steps such as resizing and grey scaling have been applied to the images in the paper.

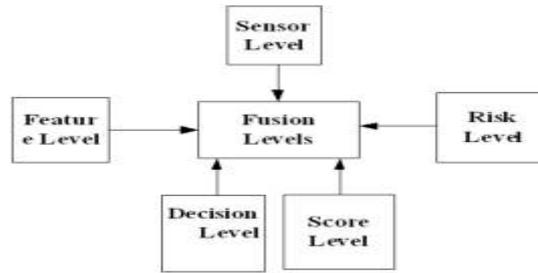


Fig. 4.5: Fusion Levels

Table 5.1: Performance of Algorithms

Algorithm	Performance Metrics		
	Face CNN	Finger CNN	Iris CNN
Accuracy	93.01	93.17	93.2
Precision	88.888	88.235	88.235
Recall	92.592	90.909	92.181
F-Score	92.0	88.235	90.909

**4.4.5. Fusion.** The fundamental module in the multimodal individual verification framework is the biometrics block. One generally utilized way to deal with the development of biometrics blocks is joining individual data from various biometric features. There are various approaches to joining biometric highlights sensor level fusion, feature level fusion, decision level fusion, score level fusion, rank level fusion as depicted in Fig 4.5 One type of data fusion is decision fusion, in which the decisions of multiple classifiers decisions are combined into a single decision about the activity that took place in five levels of fusion in multi-modal biometric systems.

*Sensor level.* where the raw data from the sensor are combined.

*Feature level.* A single feature set is created at this level by combining features extracted from each user biometric process.

*Level score:* where the final decision is made by combining match scores from different matches, which show how similar the input and stored templates are to one another.

*Rank Level.* Each character is assigned a rank by each biometric subsystem, and the ranks from the subsystems are combined to create an additional rank for identity.

*Decision Level.* By combining the result of each biometric subsystem, the final recognition decision is made.

In brief, feature-level fusion integrates biometric data or feature vectors, sensor-level fusion combines information from sensors, rank-level fusion considers the ranking of users across individual modalities, score-level fusion amalgamates similarity or matching scores from individual modalities, and decision-level fusion harmonizes the ultimate binary decisions reached by individual modalities.

We have employed decision-level fusion for this paper.

**5. Results.** Table 5.1 shows the Accuracy, Precision, Recall, and F-SCORE of face, fingerprint, and iris CNN algorithms.

The graphical representation of Table 5.1 is shown in Fig 5.1. The figure reveals the performance of the CNN algorithm on the face, fingerprint, and iris.

In the current system, the prevailing approach predominantly relies on unimodal models, with only a limited representation of multi-modal systems. Specifically, an unimodal face recognition attendance monitoring system [30] has been devised, achieving an accuracy of 87% through the utilization of the MultiTask Cascaded Convolutional Neural Network (MTCNN) and FaceNet.

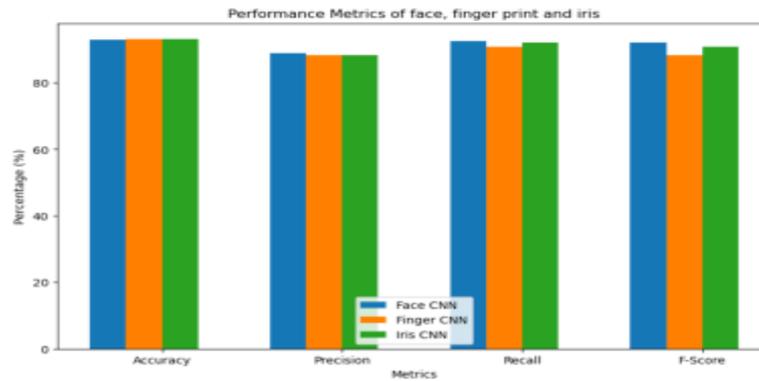


Fig. 5.1: Graphical representation of the performance of all algorithms

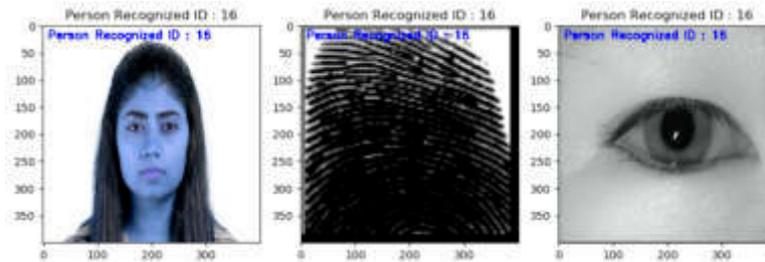


Fig. 5.2: User Prediction

In contrast, another model has adopted a multi-modal approach by integrating both facial and fingerprint biometrics [31], delivering an impressive accuracy of 96.54%. This feat was accomplished by employing the LBPH algorithm for facial recognition and the FLANN algorithm for fingerprint analysis[33]. Additionally, a separate multi-modal system, combining facial and iris recognition techniques through the deployment of the CNN algorithm, has achieved an impressive accuracy rate of 98.33% [34].

In the proposed system, we introduce a novel approach by implementing a decision-level fusion technique across three distinct biometric modalities. This represents a pioneering feature of our system. It's noteworthy that the system's accuracy can be further enhanced through the acquisition of more data and comprehensive training in future applications.

Fig 5.2 illustrates the outcome of user ID prediction achieved by inputting multimodal biometric data such as face, fingerprint, and iris.

**6. Conclusion.** The sole reason for user authentication is to guarantee that the right individuals have access to the right assets. The first step in protecting data and devices is verifying a user's identity. This paper proposes a robust convolutional neural network-based multimodal verification system to improve the current person verification system's recognition accuracy. Iris, face, and fingerprints are used in the system. Non-universality, spoof attacks, and distinctiveness are just a few of the issues that single-modal biometric systems must deal with. Implementing multimodal biometric systems, which combine multiple biometric features to overcome the difficulties of a unimodal system, can alleviate some of these limitations. The findings indicate that multimodal systems are superior to anyone biometric trait in terms of accuracy. They combine strong security with a high level of convenience.

## REFERENCES

- [1] Elhoseny, Mohamed, Ahmed Elkhateb, Ahmed Sahlol, and Aboul Ella Hassanien. "Multimodal biometric personal identification and verification." *Advances in Soft Computing and Machine Learning in Image Processing* (2018): 249-276.
- [2] Chanukya, Padira SVVN, and T. K. Thivakaran. "Multimodal biometric cryptosystem for human authentication using fingerprint and ear." *Multimedia Tools and Applications* 79 (2020): 659-673.
- [3] Mustafa, Ahmed Shamil, Aymen Jalil Abdulelah, and Abdullah Khalid Ahmed. "Multimodal biometric system iris and fingerprint recognition based on fusion technique." *International Journal of Advanced Science and Technology* 29, no. 03 (2020): 7423-7432.
- [4] Joseph, Teena, S. A. Kalaiselvan, S. U. Aswathy, R. Radhakrishnan, and A. R. Shamna. "RETRACTED ARTICLE: A multimodal biometric authentication scheme based on feature fusion for improving security in cloud environment." *Journal of Ambient Intelligence and Humanized Computing* 12, no. 6 (2021): 6141-6149.
- [5] Leghari, Mehwish, Shahzad Memon, Lachhman Das Dhomeja, Akhtar Hussain Jalbani, and Asghar Ali Chandio. "Deep feature fusion of fingerprint and online signature for multimodal biometrics." *Computers* 10, no. 2 (2021): 21.
- [6] Purohit, Himanshu, and Pawan K. Ajmera. "Optimal feature level fusion for secured human authentication in multimodal biometric system." *Machine Vision and Applications* 32 (2021): 1-12.
- [7] Wang, Yunhong, Tieniu Tan, and Anil K. Jain. "Combining face and iris biometrics for identity verification." In *Audio-and Video-Based Biometric Person Authentication: 4th International Conference, AVBPA 2003 Guildford, UK, June 9–11, 2003 Proceedings* 4, pp. 805-813. Springer Berlin Heidelberg, 2003.
- [8] Maithili, K., V. Vinothkumar, and P. Latha. "Analyzing the security mechanisms to prevent unauthorized access in cloud and network security." *Journal of Computational and Theoretical Nanoscience* 15, no. 6-7 (2018): 2059-2063.
- [9] Kirby, Michael, and Lawrence Sirovich. "Application of the Karhunen-Loeve procedure for the characterization of human faces." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 12, no. 1 (1990): 103-108.
- [10] Hong, Lin, and Anil Jain. "Integrating faces and fingerprints for personal identification." *IEEE transactions on pattern analysis and machine intelligence* 20, no. 12 (1998): 1295-1307.
- [11] Snelick, Robert, Umut Uludag, Alan Mink, Mike Indovina, and Anil Jain. "Large-scale evaluation of multimodal biometric authentication using state-of-the-art systems." *IEEE transactions on pattern analysis and machine intelligence* 27, no. 3 (2005): 450-455.
- [12] Ribaric, Slobodan, and Ivan Fratric. "A biometric identification system based on eigenpalm and eigenfinger features." *IEEE transactions on pattern analysis and machine intelligence* 27, no. 11 (2005): 1698-1709.
- [13] Monwar, Md Maruf, and Marina L. Gavrilova. "Multimodal biometric system using rank-level fusion approach." *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 39, no. 4 (2009): 867-878.
- [14] Nagajyothi, Grande, et al. "High-Speed Low Area 2D FIR Filter Using Vedic Multiplier." *Proceedings of Third International Conference on Advances in Computer Engineering and Communication Systems: ICACECS 2022*. Singapore: Springer Nature Singapore, 2023
- [15] Ma, Debbie S., Joshua Correll, and Bernd Wittenbrink. "The Chicago face database: A free stimulus set of faces and norming data." *Behavior research methods* 47 (2015): 1122-1135.
- [16] Ma, Debbie S., Justin Kantner, and Bernd Wittenbrink. "Chicago face database: Multiracial expansion." *Behavior Research Methods* 53 (2021): 1289-1300.
- [17] Lakshmi, Anjana, Bernd Wittenbrink, Joshua Correll, and Debbie S. Ma. "The India Face Set: International and cultural boundaries impact face impressions and perceptions of category membership." *Frontiers in psychology* 12 (2021): 627678.
- [18] C. Sarada, V. Dattatreya and K. V. Lakshmi, "Deep Learning based Breast Image Classification Study for Cancer Detection," 2023 IEEE International Conference on Integrated Circuits and Communication Systems (ICICACS), Raichur, India, 2023, pp. 01-08
- [19] Shehu, Yahaya Isah, Ariel Ruiz-Garcia, Vasile Palade, and Anne James. "Sokoto coventry fingerprint dataset." *arXiv preprint arXiv:1807.10609* (2018).
- [20] C. Ch.Sarada, K. V. . Lakshmi, and M. . Padmavathamma, "MLO Mammogram Pectoral Masking with Ensemble of MSER and Slope Edge Detection and Extensive Pre-Processing", *IJRITCC*, vol. 11, no. 3, pp. 135–144, Apr. 2023.
- [21] Tripathi, Shashank, Jay Murgi, Kalpana Rai3 Sneha Soni, and Rajdeep Singh. "A literature survey on multi model bio-metric system." *J. Comput. Technol* 10 (2021): 1-5.
- [22] Wang, Zijie J., Robert Turko, Omar Shaikh, Haekyu Park, Nilaksh Das, Fred Hohman, Minsuk Kahng, and Duen Horng Polo Chau. "CNN explainer: learning convolutional neural networks with interactive visualization." *IEEE Transactions on Visualization and Computer Graphics* 27, no. 2 (2020): 1396-1406.
- [23] Villavisanis, Dillan F., Clifford I. Workman, Daniel Y. Cho, Zachary D. Zapatero, Connor S. Wagner, Jessica D. Blum, Scott P. Bartlett, Jordan W. Swanson, Anjan Chatterjee, and Jesse A. Taylor. "Associations of facial proportionality, attractiveness, and character traits." *Journal of Craniofacial Surgery* 33, no. 5 (2022): 1431-1435.
- [24] Saboune, Farah Mohamad Farouk. "Virtual Reality in Social media marketing will be the new model of advertising and monetization." In *2022 Ninth International Conference on Social Networks Analysis, Management and Security (SNAMS)*, pp. 1-7. IEEE, 2022.
- [25] Cooney, Ciaran, Raffaella Folli, and Damien Coyle. "A bimodal deep learning architecture for EEG-fNIRS decoding of overt and imagined speech." *IEEE Transactions on Biomedical Engineering* 69, no. 6 (2021): 1983-1994.
- [26] Muhammad, Khan, Jamil Ahmad, Zhihan Lv, Paolo Bellavista, Po Yang, and Sung Wook Baik. "Efficient deep CNN-based fire detection and localization in video surveillance applications." *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 49, no. 7 (2018): 1419-1434.
- [27] Alaskar, Haya, Abir Hussain, Nourah Al-Aseem, Panos Liatsis, and Dhiya Al-Jumeily. "Application of convolutional neural networks for automated ulcer detection in wireless capsule endoscopy images." *Sensors* 19, no. 6 (2019): 1265.
- [28] Vinoth Kumar, V., T. Karthikeyan, P. V. Praveen Sundar, G. Magesh, and J. M. Balajee. "A quantum approach in lifi security

- using quantum key distribution." *International Journal of Advanced Science and Technology* 29 (2020): 2345-2354.
- [29] Mehmood, Asif, and Bishwas Mishra. "A Survey on Various Unimodal Biometric Techniques." *Sparklinglight Transactions on Artificial Intelligence and Quantum Computing (STAIQC)* 1, no. 1 (2021): 23-35.
- [30] Sumalatha, Veluru. "A Study of Fingerprint Patterns in Type II Diabetes." PhD diss., Rajiv Gandhi University of Health Sciences (India), 2017.
- [31] Phung, Van Hiep, and Eun Joo Rhee. "A high-accuracy model average ensemble of convolutional neural networks for classification of cloud image patches on small datasets." *Applied Sciences* 9, no. 21 (2019): 4500.
- [32] NagaJyothi, Grande, and Sriadibhatla SriDevi. "Distributed arithmetic architectures for fir filters-a comparative review." In 2017 International conference on wireless communications, signal processing and networking (WiSPNET), pp. 2684-2690. IEEE, 2017.
- [33] Rukhiran, Meennapa, Sethapong Wong-In, and Paniti Netinant. "IoT-Based Biometric Recognition Systems in Education for Identity Verification Services: Quality Assessment Approach." *IEEE Access* 11 (2023): 22767-22787.
- [34] Kumar, V. Vinoth, KM Karthick Raghunath, N. Rajesh, Muthukumaran Venkatesan, Rose Bindu Joseph, and N. Thillaiarasu. "Paddy plant disease recognition, risk analysis, and classification using deep convolution neuro-fuzzy network." *Journal of Mobile Multimedia* (2022): 325-348.
- [35] S. T. Ahmed, V. V. Kumar, and J. Kim, "AITel: eHealth Augmented-Intelligence-Based Telemedicine Resource Recommendation Framework for IoT Devices in Smart Cities." *IEEE Internet of Things*, vol. 10, no. 21, pp. 18461-18468, 1 Nov, 2023, .

*Edited by:* Polinpapilinho Katina

*Special issue on:* Scalable Dew Computing for Future Generation IoT systems

*Received:* Aug 19, 2023

*Accepted:* Nov 9, 2023



## RADIOGENOMICS IN ONCOLOGY: A COMPREHENSIVE STUDY OF VARIOUS ONCOLOGICAL DISORDERS

SOWMYA V L<sup>\*</sup>, BHARATHI MALAKREDDY A<sup>†</sup> AND SANTHI NATARAJAN <sup>‡</sup>

**Abstract.** The emergence of artificial intelligence in the digital era has brought about a significant transformation in the field of clinical decision support systems. The advent of technological advancements has led to the development of novel data-driven analytical algorithms, hence greatly augmenting human capacity to process information. The field of cancer radiogenomics presents a promising area within the realm of precision medicine. The objective of our research is to enhance our understanding of the genetic factors that contribute to the formation of tumors. This will be achieved by integrating extensive radiomics features extracted from medical imaging, genetic data obtained from clinical-epidemiological sources, and insights derived from high-throughput sequencing using mathematical modelling techniques. The aim of integrating radiomics and genomes is to gain a deeper understanding of the complex mechanisms behind cancer growth. The primary aim is to develop novel, empirically supported methodologies for the identification, prediction, and individualized therapeutic strategies for cancer, utilizing the acquired understanding. This comprehensive review aims to provide an overview of the existing body of research on the applications of radiogenomics, with a specific focus on solid malignancies. Additionally, we will examine the barriers that are now preventing the widespread integration of radiomics into therapeutic contexts.

**Key words:** Quantitative Imaging, Radiogenomics, Radiomics, Genomics, radiomics features, Machine Learning.

**1. Introduction.** Tissue biopsy remains the metric for cancer diagnosis but presents significant limitations its invasiveness, cost, and impracticality for serial monitoring make it less than ideal. In modern oncology, medical imaging serves as a fundamental aid for biopsy guidance, enhancing tumor localization and tissue sampling accuracy [1]. Imaging markers can be qualitative, quantitative, or numerically measurable attributes. The latter are the focus of radiomics, a computational approach that extracts numerous features from imaging modalities such as MRI, CT, and PET. Radiomics transform these features into actionable data, revealing patterns and correlations not immediately apparent through visual inspection. The technique offers significant promise for non-invasive disease diagnosis, prognosis, and treatment planning, thereby paving the way for personalized healthcare [2].

Genomics has revolutionized our understanding of the genetic basis of diseases, offering insights valuable for precision medicine, screening, and diagnosis. However, a significant knowledge gap persists between tissue-level imaging data and genomic information, often resulting in either over-treatment or under-treatment due to the lack of comprehensive biological markers [3].

The field of radiogenomics has recently emerged as an interdisciplinary intersection, combining the methodologies of radiomics, which involves the extraction of quantitative information from medical images, with genomics. This confluence allows for the elucidation of correlations between imaging phenotypes and genomic traits, providing a holistic approach to disease understanding and treatment planning. Notably, initial applications of 'Radiogenomics' were aimed at predicting genomic alterations from radiation therapy but have since evolved to encompass a broader range of diagnostic and prognostic capabilities [4, 7]. By leveraging data-rich patient cohorts, radiogenomics research has commenced linking imaging attributes such as tumor morphology to molecular phenotypes, a step pivotal for enhancing precision medicine, especially in oncological contexts."

This manuscript is structured in the following manner: Section 1 delves into the basic principles of radiomics and genomics. Section 2 details the radiogenomics workflow. Section 3 presents a comprehensive analysis and review of existing radiogenomics investigations. Section 4 presents a synthesis of radiogenomics applications

---

<sup>\*</sup>BMS Institute of Technology and Management, Visvesvaraya Technological University, Belagavi ([soumya8.v1@gmail.com](mailto:soumya8.v1@gmail.com)).

<sup>†</sup>Department of AI and ML, BMS Institute of Technology and Management, Visvesvaraya Technological University, Belagavi ([bharathi\\_m@bmsit.in](mailto:bharathi_m@bmsit.in)).

<sup>‡</sup>Department of CSE, Shiv Nadar University, Chennai ([santhinatarajan@snu.chennai.edu.in](mailto:santhinatarajan@snu.chennai.edu.in)).

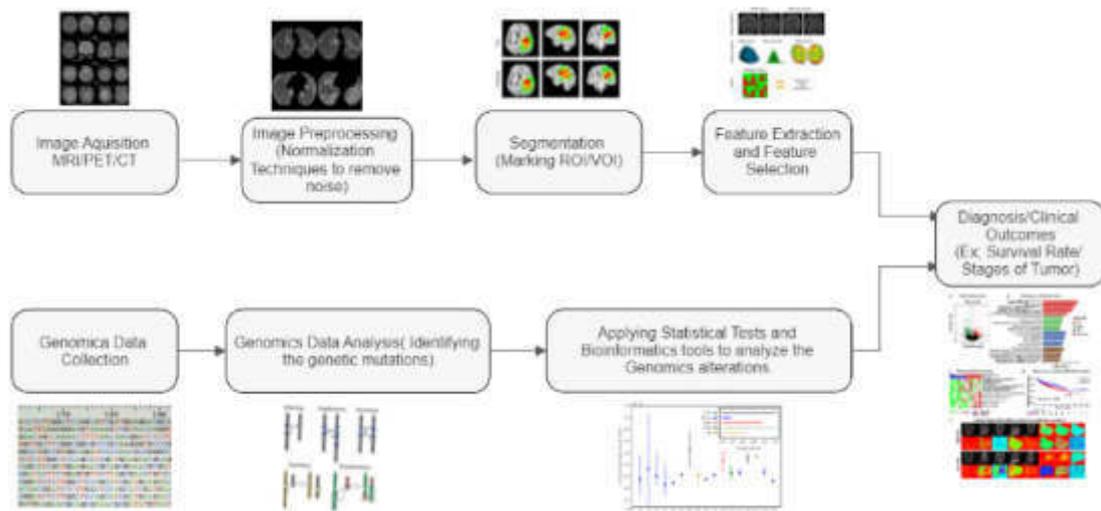


Fig. 2.1: Radiogenomics pipeline

across diverse cancer types. Section 5 highlights the constraints of such methodologies. Section 6 elaborates on the applications of radiogenomics. Finally, Section 7 concludes the paper and points towards potential avenues for future research.

**2. Overview of Radiogenomics Pipeline.** The term "Radiogenomics" refers to the relatively new field that deals with the rapid processing of radiological images and genetic information into high-dimensional information to conduct research.

Figure 2.1 represents the workflow of Radiogenomics, The radiomics and genomics pipeline involves several stages, and the implementation of radiogenomics is the integration of the data and knowledge obtained from both radiomics and genomics analyses.

**Radiomics pipeline:** It comprises the following steps: Acquisition of image, Image Pre-processing, Segmentation (ROI), Feature Extraction, Feature Selection, lastly Model Development and Validation with Disease Classification

**2.1. Radiomics pipeline.** It comprises the following steps: 1. Acquisition of image, 2. Image Pre-processing, 3. Segmentation (ROI), 4. Feature Extraction, 5. Feature Selection, lastly 6. Using machine learning creating prognostic and predictive models.

*Image Acquisition.* For the purpose of performing radiomics analysis, the original medical images are required. PET/CT/MRI have been used to aid in the detection and treatment of cancer. Images give precise details regarding the functional and structural tumor characteristics. Due to the significant advancements in medical imaging technology, radiomics has become a promising tool for addressing the challenging oncology challenges. Multi-slice CT has replaced single-slice CT in the development of medical imaging systems, enabling dynamic radiomics at various time points. CT has also been investigated to enhance the detection of tissue density [8]. Diffusion-weighted MRI can track the density, volume of tumors, and the effectiveness of cytotoxic therapy (chemotherapy that uses cytotoxic drugs).

*Image Pre-processing.* The pre-processing step is essential to gain a better high-quality image for the next analysis. photo processing steps normally consist of photograph normalization, noise removal, bias correction, interpolation and reprocessing, motion correction, and thresholding [9].

Scientific image analysis is based on the concept of picture normalization, which involves altering the range of pixel values. All records are generally implemented using standard techniques including size normalization,

histogram normalizing, and spatial normalization. Current paintings demonstrate not unusual normalization functions on more than one dataset that offer accuracy to the unique photos and enhance image segmentation.

The main function of picture noise discount is to keep the maximum vital capabilities of the photograph even as removing unimportant capabilities. Classical noise removal techniques include spatial domain filtering and variable noise removal techniques. present transform area methods together with the Fourier rework, cosine remodel, wavelet discipline technique, and sparse 3-D filtering are all advanced from the authentic spatial field approach. picture resampling is split into up-sampling and down-sampling. according to the Photo Biomarker Standardization Initiative (IBSI) [10], facts from one-of-a-kind fashions may be suitable for exceptional picture formats. movement correction is a manner to take away movement blur in a photo, which is also an important prerequisite to attaining a photo replica with the right improvement.

*Segmentation of Image.* The technique of segmentation involves dividing a picture into areas according to visual characteristics such as color, texture, density, or motion. The segment created must match the object, border, or detail area of the image [10].

Three main methods are used to segment 2-D ROI (Regions of Interest) or 3-D VOI (Volumes of Interest):

*Manual Segmentation.* Professional authors manually plot ROI / VOI or slice each image by ROI / VOI or manual segmentation or definition. processing section. Using this method, the segmentation process can be precisely controlled, as the description can take into account even the smallest changes and changes. However, it is time-consuming and leads to differences among observers.

*Semi-automatic segmentation.* Semiautomatic segmentation combines manual and automatic methods. Initially, an automated method or technique is used to provide the initial segmentation of the ROI/VOI. Users can manually update and adjust the results if needed. This approach reduces manual effort and speeds up the process while allowing users to make necessary adjustments. Examples of semi-automated processes include automated-based methods followed by manual optimization.

*Fully automatic segmentation.* Automatic segmentation refers to a fully automatic method of segmenting ROI / VOI without manual intervention using computer algorithms and machine learning. These algorithms analyze image quality such as density, texture, or shape to identify and isolate areas of interest. Automatic segmentation is generally faster and less prone to human error, but there may be differences depending on the actual algorithm used. Examples of automated segmentation techniques include clustering, edge algorithms, machine learning techniques (ML) like, CNN(convolutional neural networks), and watershed segmentation. When these methods were compared, semi-automatic segmentation was found to be the best.

ROI/VOIs delineate regions of interest in radiographs for analysis. Manual segmentation, a common approach in previous studies, doesn't involve post-processing software but is time-consuming and impractical for large datasets. Human-based delineation introduces observer-induced bias, affecting robustness due to inter-observer and intra-observer variance. Semi-automatic methods employ algorithms for ROI/VOI segmentation but may need manual correction and calibration.

Some major open-source or commercial software examples are, ITK-SNAP, 3D Slicer, ImageJ, and MITK LIFE are available as a partial partition. Automatic segmentation is based entirely on deep learning using artificial neural networks [11]. Deep learning techniques have numerous applications in automatic image recognition, particularly through Convolutional Neural Networks (CNNs). CNNs excel in computer vision tasks like image recognition and classification by extracting features directly from input data. The advantage over radiomics lies in automatic feature extraction and precise classification. However, deep learning typically requires a substantial number of medical image samples, especially when distinguishing cancerous areas from normal tissue with significant signal differences [13].

*Feature Extraction.* Segmented areas are used to extract features. Image features in medical imaging studies give crucial insights about various anatomical structures, physiological activities, and clinical conditions. Qualitative or quantitative. Radiologists and doctors may use qualitative aspects to evaluate images. Shape, margin, density, intensity, homogeneity, distribution, pattern, enhancement, calcifications, and anatomic position are described. These traits help diagnose and characterize many diseases and provide visual indications. Quantitative characteristics are calculated from imaging data and involve numerical values. Advanced image processing or automatic algorithms gain these traits.

Segmented ROIs/VOIs yield shape, statistical, and texture radiomics features. Flatness, elongation, surface

Table 2.1: Categories of quantitative radiomics features.

Category	Description	Formula
First Order Features	Basic statistical measures of pixel intensities	
	Mean	$\text{Mean} = 1/N \sum_{i=1}^N x_i$
	Median	Value at the middle position of the sorted intensity values.
	Standard deviation	$\sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{X})^2}$
	Skewness	$\frac{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{X})^3}{\left(\frac{1}{N} \sum_{i=1}^N (x_i - \bar{X})^2\right)^{\frac{3}{2}}}$
	Kurtosis	$\frac{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{X})^4}{\left(\frac{1}{N} \sum_{i=1}^N (x_i - \bar{X})^2\right)^2}$
Shape-based Features	Geometric properties of the tumor shape	
	Volume	Number of voxels within the tumor region of interest.
	Surface Area	Total surface area of the tumor region of interest.
	Compactness	$\frac{\text{Surface Area}}{\text{Volume}^{\frac{2}{3}}}$
	Sphericity	$\frac{\pi^{\frac{1}{3}} \times (6 \times \text{Volume}^{\frac{2}{3}})}{\text{Surface Area}}$
Texture-based Features	Quantify spatial patterns within the tumor	
	GLCM ( Gray-level Co-occurrence Matrix) features	Various features can be derived, such as contrast , energy, homogeneity, etc., based on the co-occurrence matrix of image intensities.
	GLRLM (Gray-level Run Length Matrix) features	Various features can be derived, such as short-run emphasis, long-run emphasis, gray-level non-uniformity, etc.
	GLSZM (Gray-level Size Zone Matrix) features	Various features can be derived, such as zone size entropy, zone percentage, zone variance, etc.
	Gray-level Dependence Matrix (GLDM) features	Various features can be derived, such as contrast, dissimilarity, homogeneity, etc., based on the dependence matrix of image intensities.
Higher-order Features	NGTDM (Neighborhood Graytone Difference Matrix) features	Various features can be derived, such as coarseness, contrast, busyness, etc., based on the neighborhood gray-tone difference matrix.
	Derived from the gray-level co-occurrence matrix (GLCM)	
	Contrast	$\sum_{i,j} P(i,j) \cdot (i,j)^2$

area, sphericity, voxel volume, and surface volume are shape characteristics. They express the lesion form, voxel intensity histogram, and texture. They can be taken directly from photos or after filters or transforms like wavelet transform. Table 2.1 lists the many quantitative radiomics features.

*Feature selection.* After features have been extracted, they are evaluated using a variety of statistical techniques to determine which ones are highly correlated from the desired results. Feature selection is a process used in radiomics to identify and select a subset of relevant features from a larger set of features extracted by the medical images. The aim is to improve the accuracy and efficiency of predictive models and reduce the dimensionality of the feature space. Table 2.2 describes the different methods for feature selection.

*Model Development and Validation with Disease Classification.* - In this pivotal stage, machine learning and statistical models are developed using integrated data, comprising radiomic, genomic, and clinical information.

- The models go through a rigorous evaluation process, which includes testing their performance using methods like cross-validation, ROC analysis, and calibration assessments. This ensures that the models work effectively and provide accurate results.

- These models are designed not only to predict patient outcomes, treatment responses, and other clinically relevant endpoints but also to excel in the accurate classification of diseases.

The utmost emphasis is placed on ensuring the reliability and generalization capabilities of these models, positioning them for invaluable clinical applications that includes disease classification, prognosis, and treatment optimization.

Table 2.2: Methods for feature selection

Type of Feature Selection Algorithm	Examples
Filter Methods	- Correlation based feature selection - Mutual information based feature selection - Chi-square feature selection
Wrapper Methods	- Recursive Feature Elimination (RFE) - Forward/Backward Feature Selection
Embedded Methods	- LASSO (Least Absolute Shrinkage and Selection Operator) - Elastic Net - Random Forest based feature selection - Gradient Boosting based feature selection
Regularization Methods	- L1 Regularization (Lasso) - L2 Regularization (Ridge)
Dimensionality Reduction Techniques	- PCA (Principal Component Analysis) - LDA (Linear Discriminant Analysis)

Table 2.3: Process of data acquisition and data preprocessing

Data Acquisition	Obtaining biological samples and their genetic data. – DNA Sequencing: Determining the order of nucleotides in a DNA sample. – Microarray Analysis: Analyzing gene expression levels for thousands of genes simultaneously. – Genotyping: Assessing specific genetic variations like SNPs.
Data Pre-processing	– Ensuring data quality, consistency, and readiness for downstream analysis. – Quality Control: Identifying and removing low – quality or erroneous data. – Data Transformation: Converting data into a suitable format for analysis if needed.

**2.2. Genomics Pipeline.** It includes the following steps genomics data acquisition and preprocessing, Genomic data analysis, Statistical analysis and modelling.

(i) Data Acquisition and Preprocessing: Collecting genomic data, such as gene sequencing or genetic profiling, from the same patients whose imaging data were used in the radiomics pipeline. Preprocessing ensures data quality and consistency shown in Table 2.3.

(ii) Genomic Data Analysis: Genomic Data Analysis is a crucial step in the genomics pipeline. It involves processing and interpreting genetic information from biological samples. Key components include variant calling, gene expression analysis, functional annotation, pathway analysis, statistical methods, and data visualization. The results contribute to scientific understanding and personalized treatment decisions.

(iii) Statistical Analysis and Modeling: Applying statistical tests and bioinformatics tools to examine the relationship among genomic alterations and clinical outcomes. This stage aims to identify genetic characteristics associated with disease development, progression, or treatment response.

**2.3. RadioGenomics Analysis.** In oncological radiogenomics, feature selection precedes the development and assessment of predictive models. These models employ selected radiogenomic attributes to forecast specific clinical outcomes in cancer patients. Model performance is rigorously evaluated through metrics such as specificity, sensitivity, and AUC-ROC, often employing cross-validation techniques. Moreover, microarray gene expression studies are integral for linking gene profiles to imaging features specific to cancer.

In radiogenomics analysis, mathematical equations are used to model and study the relationships between radiomic features and genomic data. Here are some common mathematical equations and methods used in radiogenomics.

(i) Pearson's Correlation Coefficient: The degree and direction of a linear link between two continuous

variables is measured by the Pearson's correlation coefficient, such as a radiomic feature ( $X$ ) and a genomic feature ( $Y$ ). The equation (2.1) for Pearson's correlation coefficient is,

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}} \quad (2.1)$$

where  $n$  represents the number of data points (observations),  $\sum$  denotes the summation sign, which means you need to sum up the values for all data points,  $X$  and  $Y$  are the two variables for which you are calculating the correlation coefficient (e.g., radiomic and genomic features), and  $\bar{X}$  represents the mean of variable  $X$ , and  $\bar{Y}$  represents the mean of variable  $Y$ .

(ii) Linear Regression: Linear regression models the relationship among the dependent variable (genomic feature) and one or more independent variables (radiomic features). The equation (2.2) for simple linear regression is:

$$Y = \beta_D + \beta_1 X_i + \epsilon_i \quad (2.2)$$

where  $Y$  is the genomic feature,  $\beta_0$  and  $\beta_1$  are the coefficients to be estimated, and  $X$  is the radiomic feature, and  $\epsilon$  is the error term.

(iii) Machine Learning Algorithms: Complex interactions between radiomic and genomic variables are modelled using ML algorithms including random forests, SVM (support vector machines), and neural networks. The equations for these algorithms involve specific mathematical formulations to optimize model performance and predictions.

(iv) Cox Proportional Hazards Model: In survival analysis, the Cox proportional hazards model is commonly utilized for study the link among radiomic features and patient survival or time-to-event outcomes in the presence of other covariates. The equation (2.3) for the Cox model is,

$$h(t) = h_{Q(t)} * \theta^{\sum x_i \beta_i} \quad (2.3)$$

where  $h(t)$  represents the hazard function at time,  $t$ ,  $h_0(t)$  is the baseline hazard function,  $\beta$  are the coefficients for radiomic and genomic features, and patient survival or time-to-event outcomes in the Presence of other covariates, and  $X_i$  represents the corresponding feature values.

These are radiogenomics analysis mathematical tools. Based on the type of issues, datasets, and goals, additional machine learning and statistical methods can be used for radiogenomics analysis.

**3. Review of Radiogenomics Studies.** During this research, an extensive survey was done on various popular public databases, including PubMed and Google Scholar, and found that more than a thousand radiomics papers are reported in the PubMed/MEDLINE database between 2015 and 2023. The objective of this research endeavor is to provide a comprehensive survey of the existing literature pertaining to radiogenomics for various oncological disorders. Incorporating a multitude of oncological disorders in a radiogenomics review is instrumental for elucidating the complexity and multifaceted nature of this interdisciplinary field. Varied malignancies are characterized by unique genomic aberrations and radiomic phenotypes, thereby disease-specific research is required. For example, glioblastomas frequently exhibit IDH1 mutations, whereas breast carcinomas may be scrutinized for HER2 and BRCA1/2 aberrations. Such investigations result in various outcomes, including the development of predictive models, the identification of new biomarkers, and advancements in personalized treatments. The review encompasses a wide range of cancer types, offering a comprehensive view of radiogenomic science. It also demonstrates the versatility of these methods across different areas of cancer research. This approach involves thorough comparisons, critical evaluations of current methods, and discussions about new research ideas and future directions.

**3.1. Brain.** Radiogenomics primarily focuses on central nervous system disorders. Table 3.1 summarizes neuro-oncology. Microarray DNA data and GBM neuroimmune histochemistry are used to identify tumor gene expression non-invasively [22]. In the study, 82 treatment-naive subjects from The Cancer Genome Atlas related to radiography to semantic imaging [31]. Inflammation, edema, and cell proliferation genes and microRNAs were targeted. PERIOSTIN production, a marker associated with the subtype of GBM (p 0.0001), fast relapse,

and poor prognosis, has been linked to quantitative (the measuring of quantities inside tumor compartments) ( $p < 0.001$ ).

Mutations in IDH1 and IDH2 genes improve survival in glioma patients. Thus, doctors must be able to predict an IDH1 mutation before medication without harming the patient. The decision fusion architecture created by Chang et al. for MRI images used data from various institutions on glioblastoma stages I–IV. Grinband et al. used a conventional CNN approach based on segmented regions of interest (ROI) instead of Li et al.'s Hybrid architecture, which uses a convolutional network trained with patches of ROI, a fisher vector module for encoding saliency maps, and a support vector machines classifier for estimating IDH1 mutation status. Inference on a multi-scale convolutional-only network produces feature maps represented by the fisher vector module, creating a compact representation like a bag of visual words. Multimodal low-grade glioma MRIs produced patches and ROIs. DenseNet architecture trained on multi-modal segmentations with 3D ROIs predicts IDH1 status.

DNA methylation error correcting MGMT condition. MGMT predicts GBM chemotherapeutic response. Prior work used a 50-layer ResNet, and a CNN trained on ROI data [44].

Low-grade gliomas with 1p19q co-deletion responded well to treatment and lived longer. A multi-kernel CNN [23] can predict 1p19q chromosomal expression in low-grade gliomas using T1- and T2-weighted MRI. The previous study [19] predicted the 1p19q deletion using the same deep architecture as IDH1 and MGMT. MRI texture can indicate a tumor's molecular subtype [13]. Classical (AUC=0.72) subgroups include pro-neural, mesenchymal, neural, and molecular. Radiomics effectiveness in identifying molecular subtypes encouraged the hunt for finer biological correlations. A field focused on predicting gene abnormalities used as biomarkers. T1 contrast and T2 FLAIR MR Image volumetric features linked with numerous well-known somatic mutations [43]. Mutations include TP53, RB1, NF1, EGFR, and PDGFRA. The first quantitative analysis used GBM "multiomics" data from TCGA, encompassing the transcriptome, proteome, and genome, and TCIA's matching images [15]. In a notable study, radiomic profiles were found to be associated with TP53, PTEN, and EGFR mutations [16]. The study employed random forest to discover multiple driver mutations (PDGFRA, EGFR, CDKN2A, PTEN, TP53, and RB1) based on multiparametric MRI characteristics, particularly from biopsy regions [17].

In studies of various brain tumors, including meningioma, radiomic features have been shown to group patients by grade (AUC=0.86; 81), phenotype (AUC=0.81; 81), or recurrence risk (AUC=0.72;  $p = 0.28$ ; 82). Neuroblastoma's genetic profile and medulloblastoma's molecular subtype were associated with semantics [20,21].

**3.2. Lung.** NSCLC (Non-Small Cell Lung Cancer) [40] accounts for 85% of lung cancer cases globally, with higher mortality rates. Thus, for early detection of lung cancer imaging (PET and CT) biomarkers are essential. Staging therapeutic options is challenging. EGFR (Epithelial Growth Factor Receptor) protein on cell surfaces is associated with malignancy. Predicting EGFR mutation profile can lead to improved and personalized treatments. Table 3.2 represents the review about the lung tumor [42].

EGFR is a cell surface protein that regulates cell growth and is commonly mutated in NSCLC. EGFR-targeted therapies like TKIs- Tyrosine Kinase Inhibitors are effective in treating EGFR-mutated NSCLC. KRAS-Kirsten Rat Sarcoma Viral Oncogene Homolog is a proto-oncogene linked to various cancers, including lung cancer, this is linked to a patient's treatment resistance. ALK- Anaplastic Lymphoma Kinase gene rearrangements are found in some patients of lung cancer, particularly non-smokers with adenocarcinoma, and are targeted by ALK inhibitors with remarkable efficacy. Understanding EGFR, KRAS, and ALK status is crucial for personalized lung cancer treatment, providing better outcomes with targeted therapies.

Radiomics is promising in predicting tumor genetic states, including ALK, EGFR, and KRAS [17]. KRAS shows prognostic value in various radiomic features. Effective therapeutic staging requires CT, PET scans, genetic, and laboratory biomarkers for personalized treatment decisions.

A CNN method was proposed, trained on CT patches/nodules marked by EGFR, offering non-invasive diagnostics, and supporting biopsy-identified mutations [29]. This study investigated deep-learning imaging characteristics related to NSCLC EGFR mutation status. A non-invasive investigation was used to confirm the mutational status determined by biopsy.

Genetic mutational status has become crucial in clinical decision-making for NSCLC since the FDA approved targeted therapy. NSCLC patients face challenging therapy decisions due to EGFR and KRAS muta-

tions' mutual exclusivity. KRAS mutation is a genetic alteration in the KRAS gene that leads to uncontrolled cell growth and is commonly associated with various cancers. Targeted therapies directed at these biomarkers have revolutionized lung cancer treatment, providing more effective and less toxic options for specific patient subgroups.

**3.3. Breast and Ovaries.** Since the 1970s, breast parenchyma has been studied for breast cancer risk assessment [31]. Recently, texture analysis distinguished low-risk wild-mutations from high-risk BRCA-mutations using AI methods [32]. Radiogenomics prediction models were further improved using Bayesian artificial neural networks and convoluted neural networks (AUC = 0.86) [34]. Interest in breast radiogenomics has expanded to include MRIs, with quantitative imaging linked to molecular subtypes [35]. Machine learning models recognized breast cancer characteristics, aiding risk assessment [32]. Ovarian cancer also benefits from molecular subtyping, guiding prognosis [37]. Radiogenomics holds promise for novel imaging markers with clinical potential [38]. Advancements in personalized cancer management are on the horizon. Table 3.3 represents the review of the Breast and Ovarian tumor.

The significance of molecular subtyping has also risen in the context of ovarian cancer. Accurate prognostic indicators are crucial for guiding clinical decisions, especially considering the high probability of recurrence in high-grade serous ovarian cancer [34]. Recognizing the value of merging subtype and survival gene expression patterns, researchers developed the "Classification of Ovarian Cancer" (CLO-VAR) prognostic model [36]. Semantic characteristics from a rare radiogenomic multicenter investigation involving 92 patients with high-grade serous ovarian cancer were found to be correlated with the CLOVAR system subtypes and progression time [38]. Epithelial ovarian cancer, with a 5-year survival rate ranging from 35% to 40%, necessitates precise patient categorization [39]. Radiogenomics provides insights into breast and ovarian cancer, offering novel imaging markers, enhancing personalized cancer management, and utilizing AI and machine learning, with promising potential for improved risk assessment and targeted treatments in cancer care [43].

**3.4. Liver and colorectal carcinoma.** Hepatocellular carcinoma (HCC) accounts for a significant proportion of early-stage liver cancer cases and is a leading cause of global cancer-related deaths. Fibroblast Growth Factor (FDFR) Receptor gene that has been implicated in liver cancer. These mutations may contribute to the development and progression of hepatocellular carcinoma, the most common type of primary liver cancer.

Radiogenomics research in HCC has evolved in two parallel lines, with one study using machine learning (ML) to accurately predict FDFR2 mutation in 89% of cases with a smaller sample size (n=33), demonstrating high specificity (94%) and sensitivity (87%). However, a preliminary investigation (n=66) utilizing semantic characteristics failed to detect any connection with the genetic mutation, emphasizing the need for improved prediction models through machine learning.

Colorectal cancer (CRC) is influenced by RAS gene family members that act as molecular "switches" controlling cell cycle proteins and transcription factors. KRAS mutations occur in 30-50% of CRC cases, while NRAS mutations are present in 3-5% of cases, and both are typically considered mutually exclusive. RAS mutations are associated with increased angiogenesis, cell proliferation, and metastatic potential and are used as prognostic biomarkers in the clinic, indicating EGFR antibody resistance, like NSCLC.

Various imaging techniques have been explored to develop predictive indicators for KRAS mutational status in CRC. Traditional radiomics revealed a link between KRAS mutations and skewness on CT images (p=0.02) [29]. More advanced machine learning classifiers have identified radiomics signals that predict NRAS (AUC is 0.686), BRAF (AUC is 0.857), and KRAS (AUC is 0.829) mutations with varying degrees of success. However, the lack of standardization across different tumor types hinders contrast of radiomic features associated with the similar mutation, emphasizing the need for tumor-independent radiogenomic characteristics using large datasets and advanced classification techniques.

Text-based strategies were developed to predict KRAS mutational status using descriptive language from radiology reports [26]. The trained classifier parsed radiology reports for mutant and wild-type samples, identifying specific words used more frequently for each category. The study found that KRAS-mutant tumors were often described as "few," "discrete," and "[no] recurring," while wild-type tumors were more frequently described as "multitude," "lobed," and "frequent." Further research in radiogenomics holds promise for improving cancer diagnosis and personalized treatment approaches [27]. Table 3.4 presents the review about the Liver and colorectal carcinoma.

Table 3.1: Overview of Radiogenomics methods for Brain tumor

Author	Type of Lesion	Study of Gene	Type of analysis	Image Type	Software for feature extraction	Features Identified	Methods	Results
Ziinn et. al. [13]	Gliomas	PERIOSTIN	Correlative analysis	T1 Contrast	Pyradiomics	Features of FLAIR volumes	Decision Trees	Shorter time to disease progression and Decreased survival (P<0.001)
Yang et. al. [15]	Gliomas	TP53	Correlative analysis	MRI-FLAIR	Matlab	Texture based features	Random Forest	Survival status is 0.72.
Czarnek et. al. [16]	Gliomas	POSTN	Predictive analysis	Flair MRI	Fourier Descriptor algorithm	Shape-based features	Machine Learning	Better Performance of the classification model with AUC > 0.5
Mazuriowski et al.[17]	Gliomas	EGFR, PIK3R1	Correlative analysis	MRI	matlab	Shape-based features	Machine Learning	Patient Survival rate (p=0.006)
Beig et. al. [23]	Gliomas	IDH1	Predictive analysis	T1 and T2 weighted	Matlab	First-order, shape-based features	Machine Learning	Patient Survival rate (p=0.003)
Rathore et. al. [24]	Gliomas	PIK3CA	Correlative analysis	T1 weighted and T2 weighted MRI	CapTk	Texture based features	K-Means Clustering	Categories of PTEN, TP53, EGFR genotype.
Hassan et. al. [18]	Gliomas	EGFR, PIK3R1	Correlative analysis	T1 weighted MRI	In-house radiomics pipeline	Texture based features	DNN	Patient Survival rate (p < 0.001)

**3.5. Prostate and Renal Cell Carcinoma.** Given that the clinical outcome of prostate cancer is connected closely to a primary tumor suppressor gene, radiogenomics, PTEN, has an effective promise in such cases. Table 3.5 represents the review of the prostate and renal cell carcinoma. In Prostate cancer loss of PTEN is linked to increased mortality and clinically aggressive phenotype. Where multi-parametric MR scans fail in yielding any Correlated/predictive features [32], contrast uptake on T2-weighted image intensity skewness ( $p < 0.1$ ), and DCE-MRI ( $p < 0.01$ ) is correlated with PTEN expression [30].

Radiogenomics studies were conducted on prostate cancer patients receiving MR-guided biopsies using a unique approach. Since the location of the biopsy was detected via a scan and the ROI was used for extracting radiomics features, a precise radiomics biological link could be established. This method was used to identify radiomics characteristics linked to predictive biomarkers [33,34].

The diagnosis of renal cell carcinoma (RCC) is now occurring at earlier stages, leading to more effective treatment options, primarily driven by the widespread adoption of colon cancer interventions, to reduce postoperative recurrence. The function in hypoxia signaling, Von Hippel Lindau (VHL) mutational status is frequently employed in the clinic as both a prognostic and predictive biomarker for RCC. It was shown that VHL mutations were significantly linked with nodular tumor enhancement ( $p = 0.020$ ), distinct tumor margins ( $p = 0.013$ ), and clear presence of intertumoral vascularity ( $p = 0.018$ ). Table 3.5 demonstrates that BAP1, PBRM1, and

Table 3.2: Overview of radiogenomics methods for lung tumor

Author	Type of Lesion	Study of Gene	Type of analysis	Image Type	Software for feature extraction	Features Extracted	Methods	Results
Gevaert et al. [12]	Lung tumor	EGFR	Predictive analysis	PET/CT	Matlab	Shape, edge, texturebased features	Linear regression	Better Performance of the classification model with AUC = 0.086
Sorensnet al.[13]	Lung tumor	EGFR	Predictive analysis	PET/CT	In-house computer algorithms	Edge, texture, shape based features	Classical Radiomics	EGFR mutation analysis with p = 0.05
Zhou et al. [14].	Lung tumor	KRAS	Correlative analysis	CT	matlab	Shape, Texture based features	Machine Learning	Association between phenotype and genotype
Moon et al. [15]	Lung tumor	KRAS	Correlative analysis	CT/PET	matlab	First order characteristics	Deep Learning	Pateint survival rate with progression p < 0.05
Kim et al. [16]	Lung tumor	EGFR	Correlative analysis	CT	In-line computer algorithms	Shape based features	Classical radiomics	EGFR mutation analysis with p = 0.04

VHL mutations could be detected with an accuracy of 0.75 with classifiers that use machine learning techniques validated on local datasets and tested on TCGA patients.

Radiogenomic research has focused on BAP1 mutation with VHL since it was demonstrated to be a substantial unfavorable prognostic indicator for patients with renal cell carcinoma, particularly when combined with a concurrent loss of PBRM1. To forecast the mutational status, Vikram et al, collected quantitative information from 78 cases renal cell carcinoma from the TCGA (pre-contrast AUC = 0.78) and discovered that BAP1-mutated RCCs tend to display the tumor margins (p = 0.002), CT renal vein invasion (p = 0.046) [36], and higher pathological Fuhrman grade score. Beyond specific genes, epigenetic correlations between DNA methylation in RCC and CT radiomic characteristics were also found.

#### 4. Comprehensive insights from literature review: radiogenomic applications in diverse oncological conditions.

1. Breast Cancer: Radiogenomics shows promise in differentiating between BRCAmutated highrisk and lowrisk wildtype breast cancers. Advances in AIbased texture analysis and machine learning have significantly improved predictive modeling, facilitating personalized treatment.
2. Liver Cancer (Hepatocellular Carcinoma): Distinctive advancements in HCC radiogenomics have emerged, particularly using machine learning to predict FDFR2 mutation status with high accuracy. This offers significant potential for the development of targeted therapies.
3. Ovarian Cancer: Radiogenomics has been particularly relevant in ovarian cancer subtyping and prognostication. Leveraging large multicenter datasets and advanced classification techniques, researchers aim to identify tumor-independent radiogenomic markers for refined patient classification and treatment selection.
4. Lung Cancer: Deep learning techniques in radiogenomics have been effectively used to predict EGFR mutations in non-small cell lung cancer, thereby facilitating targeted therapies such as EGFR tyrosine

Table 3.3: Overview of radiogenomics methods for breast and ovarian cancer

Author	Type of Lesion	Study of Gene	Type of analysis	Imaging Type	Feature extraction method	Features Identified	Methods	Results
Li et al. [21]	Breast tumor	BRCA	Predictive	Mammogram	Matlab	Entropy, Tumor size, shape based features	Machine Learning	AUC = 0.87
Li et. A1 [22]	Breast Tumor	BRCA	Predictive	mamogram	Computer algorithm	Volume, Area of tumor, size based features	Deep learning	AUC = 0.83
Grim m et.a1.,[24]	Breast Tumor	ER/PR/HER2	Correlative	MRI	In-house computer algorithms	Second-order characteristics features	SVM classifier	Luminal B subtype classification
Mazorowski et. al., [25]	Ovarian tumor	BRCA	Correlative	MRI	Radiomics tools	First order Features	Logistic regression	Classifier status =74 % , PR status =65 % HER2 =18 % .
Saha et. al [26]	Ovarian Tumor	HER2	Predictive	MRI	Radiomics tools	Tumor Texture, Area of Tumor, shape based features	Deep Learning	Luminal A AUC =0.697.

kinase inhibitors to improve patient outcomes.

5. Colorectal Cancer: Radiogenomics research has identified potential imaging markers for KRAS and NRAS mutational status, aiding in prognostic stratification and guiding treatment strategies.
6. Glioblastoma: Radiogenomics helps identify correlations with MGMT methylation status, enabling personalized treatment regimens including the use of alkylating agents.
7. Prostate Cancer: In prostate cancer, radiogenomics has identified MRI texture features correlating with PTEN gene deletion, informing risk stratification and treatment choices such as surgical intervention or active surveillance.
8. Methodological Advancements: Recent strides in integrative radiogenomics frameworks that combine traditional statistical methods with machine learning have increased both the robustness and predictive accuracy of the models.
9. Ethical and Regulatory Considerations: Issues such as patient consent, data sharing, and algorithmic transparency are gaining attention, calling for ethical guidelines and regulatory oversight for clinical adoption.
10. Clinical Translation and Validation: Despite advancements, direct clinical application remains nascent, requiring further large-scale, multi-center trials for validation and reliability assessment.

**5. Limitation of Radiogenomics.** Limitations are derived from literature review and expert consultations. Existing literature revealed gaps such as sample heterogeneity and algorithmic biases. Expert input confirmed these limitations and provided detailed insights into technical and ethical challenges. This dual approach ensures that the limitations mentioned are supported by academic foundations and validated through the opinions of current experts.

Table 3.4: Overview of radiogenomics methods for liver and colorectal carcinoma

Author	Type of Lesion	Study of Gene	Type of analysis	Image type	Feature extraction method	Features Extracted	Methods	Results
Kuo et al. [29]	Liver Tumor	TP53	Correlative	PET/CT	CapTk	Grey Level Mean, Maximum, Shape, etc.	Machine Learning	Survival rate $p < 0.05$
West et al. [28]	Liver Tumor	Metastasis TP53	Predictive	CT	Matlab	Shape, Tumox size, area of the tumot, Volumetric etc,	Classical Radiomics	Confirmed TP53 AUC is 86.61 %, Specificity is 92.31 %, Sensitivity is 92.9 %.
Lubner et al. [30]	Colorectal Tumor	KRAS	Correlative	PET	Radiomics tools	First order characteristics	Machine Learning	KRAS mutations were negatively associated with Skewness ( $P=0.02$ ).
Lovinfosse et al. [31]	Colorectal Tumor	KRAS	Correlative	CT	Matlab tools	Grey Level Mean Maximum, Shape etc.	Deep Learning	Eighty-three patients had RAS mutations: 9 NRAS, 74 KRAS and 68 patients had no changes.

Limited sample sizes: Some radiogenomics studies may have small sample sizes, which can limit the generalizability of the findings.

Data heterogeneity: Variability in imaging protocols and equipment across different institutions can introduce heterogeneity in radiomic data, affecting the consistency of results.

Retrospective data: Many radiogenomics studies rely on retrospective data, which may lead to selection bias and other limitations in study design.

Standardization challenges: Lack of standardized radiomic features and methodologies can make it difficult to compare results across different studies.

Overfitting: Complex machine learning algorithms used in radiogenomics may lead to overfitting, where the model performs well on training data but poorly on new data.

**6. Application of Radiogenomics.** Radiogenomics, the integrated study of radiomic and genomic data, has been increasingly recognized for its potential in oncology. The literature reveals extensive applications of radiogenomics in predicting treatment responses, distinguishing aggressive phenotypes, identifying genomic alterations, and determining tumor heterogeneity. Moreover, its utility extends to predicting patient prognosis, facilitating the understanding of potential metastatic risks, and aiding in the selection of suitable therapeutic strategies. These advancements highlight the transformative potential of radiogenomics in tailoring precision medicine approaches, underscoring its significant role in enhancing diagnostic and therapeutic decisions.

*Survival Analysis.* Conceptual Framework: Survival analysis in radiogenomics seeks to correlate radiomic features extracted from medical images with genomic profiles to predict patient outcomes, such as overall survival or disease-free survival.

Table 3.5: Overview Of Radiogenomics Methods For Prostate Cancer And Renal Cell Carcinoma

Author	Type of Lesion	Study of Gene	Type of analysis	Imaging type	Feature extraction method	Features Extracted	Methods	Results
Vander Weele et al. [32]	Prostrate tumor	PTEN	Correlative	MRI	In-house computer algorithms	Shape, Texture, edge etc.	Machine Learning	Feature values with expressions less than 0.25 and interquartile ranges (IQRs) less than 0.5 were filtered for significant representation
Mc Cann et al. [33]	Prostrate tumor	PTEN	Predictive	MRI	CapTk	Edge, texture, shape.	Classical Radiomics	Binary classification of prostate cancer.
Stoyanova et al. [33, 34]	Prostrate tumor	General gene expression	Correlative	MRI	Matlab	Shape, Texture	Machine Learning	Identification of prostrate tumor
Shinagare et al. [36]	Renal tumor	BAP1	Correlative	CT	In-house computer algorithms	First order characteristics	Deep Learning	Correlation between BAP1 and features for Renal tumor.
Karlo et al. [37]	Renal tumor	PBRM1, VHL	Correlative	MRI	In-house computer algorithms	Nodular tumor	Classical radiomics	Correlation between PBRM1 and features for Renal tumor.

A commonly used model is the Cox Proportional Hazards Model is calculated in equation (6.1),

$$h(t | X) = h_0(t) \exp(\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_y x_y) \quad (6.1)$$

Here,  $h(t | X)$  is the hazard function dependent on time  $t$  and covariates  $X$ , and  $h_0(t)$  is the baseline hazard.

Figure 6.1 depicts a Kaplan-Meier survival curve, which is commonly used in medical research to represent the fraction of patients living for a certain amount of time after treatment. Kaplan-Meier survival curve, extracted from a comprehensive literature review, showcases the survival probabilities of high-risk and low-risk patient cohorts over time. Its strength lies in adeptly representing censored data and enabling direct comparisons between groups. The curve's trajectory reveals significant survival differences between the cohorts, supported by the provided p-value. Kaplan-Meier's visual clarity ensures its widespread use in oncology research, offering immediate insights for experts and practitioners.

Figure 6.1 has been adapted to vividly illustrate survival probabilities over a 12-year span [41]. The X-axis marks time in years, while the Y-axis denotes survival probability from 0 to 1. The high-risk group, represented

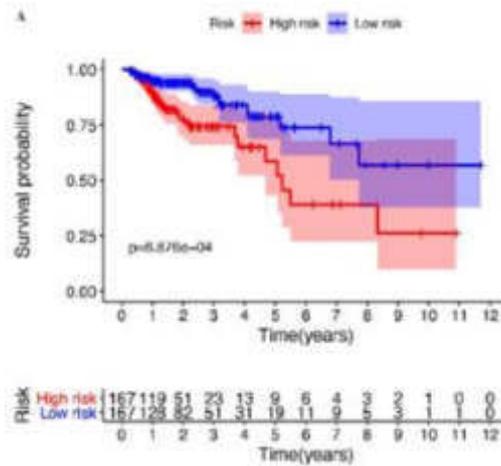


Fig. 6.1: Comparative Kaplan-Meier Survival Analysis of High-risk vs. Low-risk Patient Cohorts

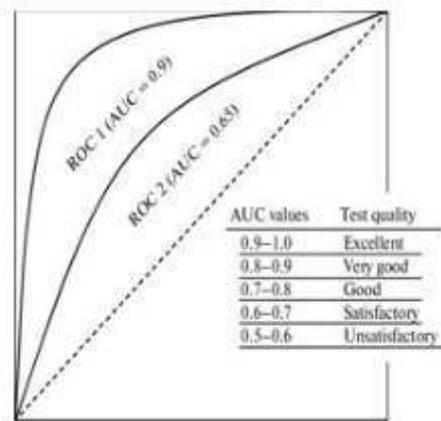


Fig. 6.2: Evaluation of Diagnostic Test Performance: ROC Curves and AUC Value Interpretation for Metastasis Prediction

by the red curve, showcases a steeper decline in survival compared to the blue curve of the low-risk group. At year 0, both cohorts start near a 1.0 survival probability, but by year 12, the high-risk group approaches 0.25 while the low risk remains above 0.5. Cross marks indicate censored data points, where exact survival times are unknown. The significant p-value of 8.876e-04 emphasizes the meaningful difference between these two groups. The risk table below provides a numerical breakdown of patients remaining in each cohort over the years.

*Predicting Metastasis.* Conceptual Framework: Radiogenomics can predict the likelihood of metastasis by linking image-derived radiomic features with genomic markers known to be associated with metastatic spread.

Logistic regression can model this binary outcome is calculated as in equation (6.2),

$$\log\left(\frac{g}{1-\eta}\right) = \beta_0 + \beta_1x_1 + \beta_2x_2 \dots \dots \dots \tag{6.2}$$

Here, p represents the probability of metastasis occurring

The depicted diagram, Figure 6.2, showcases two Receiver Operating Characteristic (ROC) curves, instru-

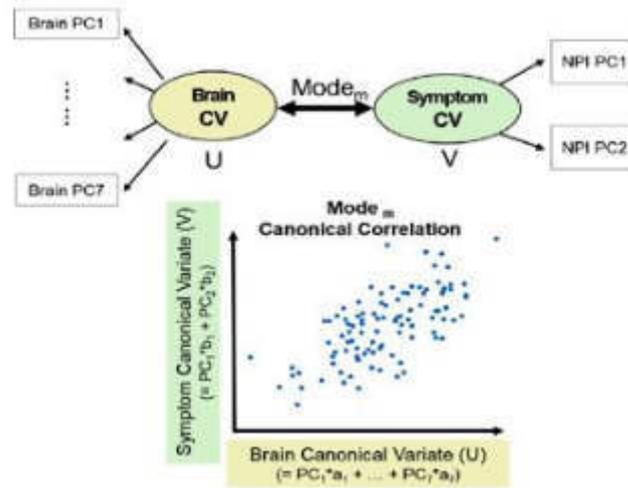


Fig. 6.3: Canonical Correlation Analysis between Brain Principal Components and Symptom Severity

mental in evaluating the performance of diagnostic tests, especially in predicting conditions like metastasis in oncology [34]. The ROC curve, as illustrated in Figure 6.2, is a graphical representation of the true positive rate (sensitivity) against the false positive rate (1-specificity) for various threshold values. A perfect diagnostic test would result in a curve passing through the top-left corner, indicating 100% sensitivity and 100% specificity.

Within Figure 6.2, two ROC curves are displayed: ROC1 with an Area Under the Curve (AUC) of 0.9 and ROC2 with an AUC of 0.65. AUC is a metric capturing the overall performance of a diagnostic test, ranging from 0.5 (no discrimination) to 1.0 (perfect discrimination). Here, ROC1, boasting an AUC of 0.9, denotes "Very Good" test quality, implying a high accuracy in predicting metastasis. In contrast, ROC2, also presented in Figure 6.2, with an AUC of 0.65 signifies just a "Satisfactory" test quality. The accompanying table within Figure 6.2 provides a categorical assessment of AUC values, guiding interpretations of test efficacies.

*Integrative Approaches.* For simultaneous analysis of radiomic and genomic data, Canonical Correlation Analysis (CCA) is often used shown in equation (6.3)

$$\text{Maximize } u^T X v \quad (6.3)$$

Here  $u$  and  $v$  are weight vectors for the radiomic and genomic datasets, and  $X$  is the correlation matrix.

Figure 6.3 depicts a graphical representation of the canonical correlation analysis, a statistical method utilized to ascertain the relationship between two sets of variables [39]. In this case, the two sets are the principal components (PCs) derived from brain data (Brain PC1 to Brain PC7) and the principal components from symptom data (NPI PC1 and NPI PC2). The purpose of canonical correlation is to find pairs of linear combinations, one from each set, that have maximum correlation with each other.

In the depicted Figure 6.3, these linear combinations are referred to as "Brain CV" (canonical variate U) and "Symptom CV" (canonical variate V). The scatter plot showcases the canonical correlation for a specific mode (Mode  $m$ ), plotting the relationship between the Brain Canonical Variate and the Symptom Canonical Variate. As indicated by the equations at the bottom, each canonical variate is a linear combination of its respective PCs, weighted by certain coefficients ( $a$ 's for the brain). Visualization aids in discerning the strength and nature of the relationship between the brain's structural features and symptom severity, offering invaluable insights into potential neurological underpinnings of the presented symptoms.

**7. Conclusion and Future work.** In this comprehensive review, the salient advancements, and challenges in the domain of radiogenomics research are highlighted, especially in the context of various oncological disorders. Radiogenomics, grounded in earlier research, is still an emerging field. Despite the existing challenges, significant progress has been made in understanding and addressing various tumor types. The advent

of deep learning and advanced artificial intelligence techniques in medicine shows potential in navigating the current barriers in the practical utilization of radiogenomics.

As the horizon unfolds, the integration of these cutting-edge technologies is expected to redefine the clinical paradigm in radiogenomics. This transformative shift calls for radiologists to adapt and actively immerse themselves in the innovations. The ongoing growth of this interdisciplinary realm will likely lead to more tailored therapeutic approaches for cancer care, emphasizing the imperative for future studies to refine and standardize these burgeoning methods.

**Acknowledgments.** This research is funded by Vision Group on Science and Technology (VGST), Government of Karnataka, India.

#### REFERENCES

- [1] Gagandeep Singh, Sunil Manjila, Nicole Sakla, Alan True, "Radiomics and radiogenomics in gliomas: a contemporary update", *British Journal of Cancer* 75:641–657, 2021.
- [2] Eleftherios trivizakis, Georgios z. papadakis, "Artificial intelligence radiogenomics for advancing precision and effectiveness in oncologic care", *INTERNATIONAL JOURNAL OF ONCOLOGY*, PMC 57(1): 43–53, 2020.
- [3] Coroller TP, Bi WL, Huynh E, Abedalthagafi M, Aizer AA, Greenwald NF, Parmar C, Narayan V, Wu WW, de Moura SMJPo, "Radiographic prediction of meningioma grade by semantic and radiomic features", *PubMed* 12 (11):e01878902017.
- [4] Yip SS, Liu Y, Parmar C, Li Q, Liu S, Qu F, Ye Z, Gillies RJ, Aerts HJ, "Associations between radiologist-defined semantic and automatically computed radiomic features in non- small cell lung cancer", *Scientific reports* 7 (1):3519, 2017.
- [5] Gillies RJ, Kinahan PE, Hricak H, "Radiomics: Images Are More than Pictures, They Are Data". *Radiology* 278 (2):563-577, 2016.
- [6] Court LE, Rao A, Krishnan S, "Radiomics in cancer diagnosis, cancer staging, and prediction of response to treatment", *Translational cancer research* 5 (4):337-339, 2019.
- [7] Rutman AM, Kuo MD, "Radiogenomics: creating a link between molecular diagnostics and diagnostic imaging", *European journal of radiology* 70 (2):232-241, 2019.
- [8] Kang J, Rancati T, Lee S, Oh JH, Kerns SL, Scott JG, Schwartz R, Kim S, Rosenstein BS, "Machine Learning and Radiogenomics: Lessons Learned and Future Directions", *Frontiers in oncology* 8:228, 2018.
- [9] Antropova N, "Deep Learning and Radiomics of Breast Cancer on DCE-MRI in Assessment of Malignancy and Response to Therapy", *The University of Chicago ProQuest Dissertations Publishing* 10784160, 2018.
- [10] Aerts HJ, "The Potential of Radiomic-Based Phenotyping in Precision Medicine: A Review", *JAMA oncology* 2(12):1636-1642, 2016.
- [11] Berger MF, Mardis ER, "The emerging clinical relevance of genomics in cancer medicine", *Nat Rev Clin Oncol* 15(6):353-365, 2018.
- [12] Zinn PO, Mahajan B, Sathyan P, Singh SK, Majumder S, Jolesz FA, Colen RR, "Radiogenomic mapping of edema/cellular invasion MRI-phenotypes in glioblastoma multiforme", *PloS One* 6(10): e25451, 2011.
- [13] Zhao L, Lee VH, Ng MK, Yan H, Bijlsma, "Molecular subtyping of cancer: current status and moving toward clinical applications", *Brief Bio inform Journal*, *PubMed* 25;20(2):572-584, 2018.
- [14] Yang D, Rao G, Martinez J, Veeraraghavan A, Rao A, "Evaluation of tumor-derived MRI-texture features for discrimination of molecular subtypes and prediction of 12-month survival status in glioblastoma", *Medical physics PubMed* 42(11):6725-35, 2015.
- [15] Czarnek NM, Clark K, Peters KB, Collins LM, Mazurowski MA, "Radiogenomics of glioblastoma: A pilot multi-institutional study to investigate a relationship between tumor shape features and tumor molecular subtype", *SPIE Medical Imaging*, 1605-7422, 2016.
- [16] Mazurowski MA, Clark K, Czarnek NM, Shamsesfandabadi P, Peters KB, Saha A (2017) Radiogenomics of lower-grade glioma: algorithmically assessed tumor shape is associated with tumor genomic subtypes and patient outcomes in a multi- institutional study with The Cancer Genome Atlas data. *Journal of neuro-oncology* 133(1):27-35, 2017.
- [17] Zinn PO, Singh SK, Kotrotsou A, Abrol S, Thomas G, Mosley J, Elakkad A, Hassan I, Kumar A, Colen RR, "Distinct Radiomic Phenotypes Define Glioblastoma TP53-PTEN-EGFR Mutational Landscape", *Neurosurgery* 64(CN Suppl 1): 203–210, 2017.
- [18] Hu LS, Ning S, Eschbacher JM, Baxter LC, Gaw N, Ranjbar S, Plasencia J, Dueck AC, Peng S, Smith KA, Nakaji P, Karis JP, Quarles CC, Wu T, Loftus JC, Jenkins RB, Sicotte H, Kollmeyer TM, O'Neill BP, Elmquist W, Hoxworth JM, Frakes D, Sarkaria J, Swanson KR, Tran NL, Li J, Mitchell, "Radiogenomics to characterize regional genetic heterogeneity in glioblastoma", *Neuro-oncology* 19(1): 128-137, 2017.
- [19] Kickingereder P, Bonekamp D, Nowosielski M, Kratz A, Sill M, Burth S, Wick A, Eidel O, Schlemmer HP, Radbruch A, Debus J, Herold-Mende C, Unterberg A, Jones D, Pfister S, Wick W, von Deimling A, Bendszus M, Capper D, "Radiogenomics of Glioblastoma: Machine Learning-based Classification of Molecular Characteristics by Using Multiparametric and Multi-regional MR Imaging Features", *Radiology PubMed* 281(3):907-918, 2016.
- [20] Li H, Giger ML, Lan L, Yuan Y, Bhooshan N, Olopade, "Effect of variable gain on computerized texture analysis on digitalized mammograms", *SPIE Medical Imaging*, 7624:76242C, 2010.

- [21] Li H, Giger ML, Huynh BQ, Antropova, "Deep learn ing in breast cancer risk assessment: evaluation of convolutional neural networks on a clinical dataset of full-field digital mammograms", *Journal of medical imaging* 4(4):041304, 2017.
- [22] Beig N, Patel J, Prasanna P, Partovi S, Varadan V, Madabhushi A, Tiwari P, "Radiogenomic analysis of hypoxia pathway reveals computerized MRI descriptors predictive of overall survival in glioblastoma", *SPIE* 8(7): 1-11,2017.
- [23] Rathore S, Akbari H, Rozycki M, Abdullah KG, Nasrallah MP, Binder ZA, Davuluri RV, Lustig RA, Dahmane N, Bilello M, O'Rourke DM, Davatzikos C, "Radiomic MRI signature reveals three distinct subtypes of glioblastoma with different clinical and molecular characteristics offering prognostic value beyond IDH1", *Scientific reports* 5087(8):1-12, 2018.
- [24] Mazurowski MA, Zhang J, Grimm LJ, Yoon SC, Silber JI, "Radiogenomic analysis of breast cancer: luminal B molecular subtype is associated with enhancement dynamics at MR imaging", *Radiology Proc SPIE Int Soc Opt Eng* 10950: 1095044, 2016.
- [25] Saha A, Harowicz MR, Grimm LJ, Kim CE, Ghate SV, Walsh R, Mazurowski MA, "A machine learning approach to radiogenomics of breast cancer: a study of 922 subjects and 529 DCE-MRI features", *British journal of cancer* 119(4):508-516, 2018.
- [26] Zhu Y, Li H, Guo W, Drukker K, Lan L, Giger ML, Ji Y, "Deciphering Genomic Underpinnings of Quantitative MRI- based Radiomic Phenotypes of Invasive Breast Carcinoma", *Scientific reports* 5:17787, 1-12, 2015.
- [27] West DL, Kotrotsou A, Niekamp AS, Idris T, Camejo DG, Mazal NJ, Cardenas NJ, Goldberg JL, Colen RR, "CT-based radi omics analysis of hepatocellular carcinoma patients to predict key genomic information", *Journal of Clinical Oncology Conference PubMed* 14(12): 2860, 2017.
- [28] Kuo MD, Gollub J, Sirlin CB, Ooi C, Chen X, "Radiogenomic analysis to identify imaging phenotypes associated with drug response gene expression programs in hepatocellular carcinoma. *Journal of vascular and interventional radiology*", *JVIR* 18(7):821-830, 2018.
- [29] Lubner MG, Stabo N, Lubner SJ, del Rio AM, Song C, Halberg RB, Pickhardt PJ, "CT textural analysis of hepatic meta- static colorectal cancer: pre-treatment tumor heterogeneity correlates with pathology and clinical outcomes", *Abdominal imaging PubMed* 40(7):2331-7, 2015.
- [30] Lovinfosse P, Koopmansch B, Lambert F, Jodogne S, Kustermans G, Hatt M, Visvikis D, Seidel L, Polus M, Albert A, Delvenne P, Hustinx R, "F-FDG PET/CT imaging in rectal cancer: relationship with the RAS mutational status", *The British journal of radiology* 89(1063): 20160212, 2016.
- [31] Vander-Weele DJ, McCann S, Fan X, Antic T, Jiang Y, Oto A, "Radiogenomics of prostate cancer: Association between quantitative multi-parametric MRI features and PTEN", *Journal of Clinical Oncology Conference* 33:126-128, 2015.
- [32] McCann SM, Jiang Y, Fan X, Wang J, Antic T, Prior F, "Quantitative Multiparametric MRI Features and PTEN Expression of Peripheral Zone Prostate Cancer: A Pilot Study", *AJR American journal of roentgenology* 206(3):559-65, 2016.
- [33] Stoyanova R, Pollack A, Takhar M, Lynne C, Parra N, Lam LL, Alshalalfa M, Buerki C, Castillo R, Jorda M, Ashab HA, Kryvenko ON, Punnen S, Parekh DJ, Abramowitz MC, Gillies RJ, Davicioni E, Erho N, Ishkanian A, "Association of multiparametric MRI quantitative imaging features with prostate cancer gene expression in MRI-targeted prostate biopsies", *Oncotarget* 7(33): 53362-53376, 2019.
- [34] Stoyanova R, Pollack A, Lynne C, Jorda M, Erho N, Lam LLC, Buerki C, Davicioni E, Ishkanian A, "Using radiogenomics to characterize MRI-guided prostate cancer biopsy heterogeneity", *Journal of Clinical Oncology Conference* 33:7\_suppl, 25, 2015.
- [35] Shinagare AB, Vikram R, Jaffe C, Akin O, Kirby J, Huang E, Freymann J, Sainani NI, Sadow CA, Bathala TK, Rubin DL, Oto A, Heller MT, Surabhi VR, Katabathina V, Silverman SG, "Radiogenomics of clear cell renal cell carcinoma: preliminary findings of The Cancer Genome Atlas-Renal Cell Carcinoma (TCGA-RCC) Imaging Research Group", *Abdominal imaging* 40(6):1684-92, 2015.
- [36] Karlo CA, Di Paolo PL, Chaim J, Hakimi AA, Ostrovnaya I, Russo P, Hricak H, Motzer R, Hsieh JJ, Akin O, "Radiogenomics of clear cell renal cell carcinoma: associations between CT imaging features and mutations", *Journal of Radiology PubMed* 270(2):464-71, 2014.
- [37] Gevaert O, Xu J, Hoang CD, Leung AN, Xu Y, Quon A, Rubin DL, Napel S, Plevritis SK, "Non-small cell lung cancer: identifying prognostic imaging biomarkers by leveraging public gene expression microarray data methods and preliminary results", *Journal of Radiology* 264(2):387-96, 2012.
- [38] Sorensen J, Erasmus Jr J, Shroff G, Rao A, Stingo F, Court L, Gold KA, Lee J, Heymach JV, Swisher S, Godoy, "Combining CT texture analysis with semantic imaging descriptions for the radiogenomic detection of EGFR and KRAS mutations in NSCLC", *Journal of Thoracic Oncology MDPI* 14(7): 1657, 2019.
- [39] Zhou M, Leung A, Echegaray S, Gentles A, Shrager JB, Jensen KC, Berry GJ, Plevritis SK, Rubin DL, Napel S, Gevaert O, "Non-Small Cell Lung Cancer Radiogenomics Map Identifies Relationships between Molecular and Imaging Phenotypes with Prognostic Implications", *Journal of Radiology PubMed* 286(1):307-315, 2018.
- [40] Moon H, Choi JH, Park JH, Byun BH, Lim I, Kim BI, Choi CW, Lim SM, "Differences of FDG PET tumor texture parameters according to the presence of epidermal growth factor receptor mutation in non-small cell lung cancer", *Journal of Nuclear Medicine Conference: Society of Nuclear Medicine and Molecular Imaging Annual Meeting, SNMMI* 31(4): 632-640, 2015.
- [41] Kim TJ, Lee CT, Theon SH, Park JS, Chung JH, "Radiologic Characteristics of Surgically Resected Non-Small Cell Lung Cancer with ALK Rearrangement or EGFR Mutations", *The Annals of Thoracic Surgery* 101(2):473-80, 2016.
- [42] Parita Oza, Paawan Sharma, and Samir Patel, "A Transfer Representation Learning Approach for Breast Cancer Diagnosis from Mammograms Using EfficientNet Models", *Scalable Computing: Practice and Experience* 23(2):51-58, 2022.
- [43] Parasa Rishi kumar, Kavya Bonthu, Boyapati Meghana, Koneru Suvarna Vani, and Prasun Chakrabarti, "Multi-class Brain Tumor Classification and Segmentation Using Hybrid Deep Learning Network (HDLN) Model", *Scalable Computing: Practice and Experience* 24(1):69-80, 2023.

- [44] Madhavi Latha Pandala, S. Periyamayagi, “Analysis, Prediction and Classification of Skin Cancer using Artificial Intelligence - A Brief Study and Review”, SCPE 24(3): 355–370, 2023.

*Edited by:* Polinapilinho Katina

*Special issue on:* Scalable Dew Computing for Future Generation IoT systems

*Received:* Aug 25, 2023

*Accepted:* Dec 4, 2023



## SENSOR BASED DANCE COHERENT ACTION GENERATION MODEL USING DEEP LEARNING FRAMEWORK

HANZHEN JIANG \*AND YINGDONG YAN †

**Abstract.** Dance Coherent Action Generation is a popular research task in recent years to generate movements and actions for computer-generated characters in a simulated environment. It is sometimes referred to as "Motion Synthesis". Motion synthesis algorithms are used to generate physically believable, visually compelling, and contextually appropriate movement using motion sensors. The Dance Coherent Action Generation Model (DCAM) is a generative framework for producing aesthetically pleasing movements using deep learning from small amounts of data. By learning an internal representation of motion dynamics, DCAM can synthesize long sequences of movements in which coherent patterns can be created through latent space interpolation. This framework provides a mechanism for varying the amplitude of the generated motion, allowing further realistic thinking and expression. The proposed model obtained 93.79% accuracy, 93.79% precision, 97.75% recall and 92.92% F1 score. DCAM exploits the balance between imitation and creativity by enabling the production of novel outputs from limited input data and can be trained in an unsupervised manner or fine-tuned with sparse supervision. Furthermore, the framework is easily extended to handle multiple layers of abstraction and can be further personalized to a particular type of movement, enabling the generation of highly individualized outputs.

**Key words:** Dance, Coherent, Action, Generation, Computer Animation, Video Games, Sensor.

**1. Introduction.** Dance is an activity that allows people to express themselves and cultivate relationships through creative and meaningful body movements. Recently, it has developed into a popular form of art and recreation. It is a great way to promote physical activity and socialization and build self-confidence. Generating coherent action in dancing is a process that involves developing knowledge of the various techniques, steps, and styles associated with a chosen dance [1]. It can be done by studying the history and culture behind a particular dance style, understanding its body movements, and learning the appropriate techniques and steps through practice. In order to generate coherent action in dancing requires analytical thinking and creative expression. It involves understanding the context in which the dance takes place, such as the atmosphere, music, and the audience. It also involves being able to recognize the dynamics and nuances of the dance and have the ability to be flexible in order to adapt to the changing environment [2]. It is essential to have an idea of the structure of the dance and how it can be manipulated to create the desired choreography. Generating coherent action in dancing also necessitates body awareness and control. It is essential to identify the correct muscles and movements used in the dance and utilize them in a planned and strategic way [3]. Finally, it involves understanding how to convey emotion and rhythm through movement and developing a good sense of coordination, timing, and balance. Generating coherent action in dancing is an intricate process that requires analytical thinking, creativity, body awareness, and control. Through practice and dedication, one can successfully generate coherent action in dancing and bring to life a unique and enjoyable experience for all. Dance has existed since the beginning of time and is constantly evolving and innovating. In modern times, the development of technology has had a significant impact on how dance is created, performed, and taught [4]. One of the most significant innovations in dance is the development of coherent action generation. This technology allows skilled dancers to generate complex movements based on a given set of goals or objectives. By breaking down sophisticated choreographic processes into discrete elements, these movements can be performed and perfected much more quickly. It has allowed professional dancers to create stunning and complex routines without spending hours creating them from scratch. Another development within the dance industry is the use of motion capture technology [5]. This technology is used to map the movements of a dancer in order to create a 3-D interpretation of their routine.

---

\*International College, Krirk University, Bangkok, 10220, Thailand ([hanzhen93@sina.com](mailto:hanzhen93@sina.com))

†International College, Krirk University, Bangkok, 10220, Thailand ([791086531@qq.com](mailto:791086531@qq.com))

It allows the choreographer to analyze and adjust the routine as needed to perfect the performance. Thanks to motion capture technology, 3-D-created routines can also be studied and replicated by others. The technology used to create virtual simulations is advancement in dance [6]. With virtual simulations, dancers can visualize different stages of their performance in a manner that were not possible before. It is allowing 3-D choreography to become commonplace, even with opener-level dancers. This technology is used to plan out routines and in the editing process once a dance has been completed. It is essential to note the increased use of smart phones and other mobile devices in dance. These devices allow dancers to access instructional videos, lessons, and other resources from almost anywhere [7]. It has also allowed them to connect with choreographers and other dancers to further their skills. The advancements in technology are revolutionizing how dancers approach and perform their art. Through coherent action generation, motion capture, virtual simulations, and mobile devices, dancers can create complex routines and take their craft to new heights.

The research and practice of dance coherent action generation is a field of study that investigates how computers can create sequences of movements to provide a dancer or a group of dancers with a logical, fluid, and aesthetically pleasing pattern of movements. Deep Learning (DL) is crucial in this research as neural networks can learn from example data and generate predictive models [8]. This research and application of DL methodologies has become increasingly popular in dance and the automated generation of coherent motion sequences. One of the most successful approaches to dance coherent action generation is using Long Short-Term Memory (LSTM) networks. LSTMs can capture short-term and long-term dependencies in the motion of a dancer. Such networks can generate coherent motion sequences based on an initial input path and influential parameters, bypassing the traditional choreography process. In addition, such networks allow a controllable degree of randomness, which affects the generated motions, enhancing the expressive qualities of the dance sequence [9]. Combined with further recurrent neural networks, this model enables a projective technique that limits the rate of change and preserves the timing characteristics of sequences. DL networks have been successfully applied to various domains of the arts, such as music and visual arts. However, issues of aesthetics and expression have yet to be considerably explored in research on computer choreography. Recent Reinforcement Learning (RL) developments have provided promising solutions for this aspect of the field. Such solutions use a reward function and principles of imitation learning to generate coherent and expressive movements [10]. As computer choreography and DL-aided motion generation mature, these technologies will become more common as they present an opportunity to reduce the time and effort necessary for the manual creation of choreography. Nevertheless, further research is needed to address some of the limitations of current DL techniques for dance and accompanying expressive qualities, such as the ability to handle highly dynamic and expressively charged dances. The main contribution of the research has the following,

- Developing and validating metrics for evaluating the quality of dance motions and generating realistic motions using data-driven methods such as motion capture data;
- Exploring different ways of combining multiple motion sources to produce more fluid and consistent motions and developing motion models capable of learning to generate novel dance motion;
- Understanding the underlying formal grammar of dance motion and modeling physical dynamics for motion synthesis;
- To design efficient motion representation and scheduling algorithms for efficient online motion generation and developing interactive systems for generating coherent dance sequences.

**2. Related Work.** Muthukumaran, V et al. [11] have discussed Motion generation based on a multi-feature fusion strategy is a technique used to generate motion based on multiple features of a body or object. It requires the fusion of pre-experimentally acquired features from various sensors, such as motion sensors, wearable biosensors, and inertial measurement units. This technique is used to generate realistic and robust motions for robotic applications and animation and gaming applications. It is also used in medical applications, rehabilitation, and training applications in sports and other disciplines.

Zhang et al. [12] have discussed Mining and applying composition knowledge of dance moves for style-concentrated dance generation is a process that utilizes artificial intelligence techniques (such as machine learning and deep learning) to analyze existing dance moves and styles from different choreographers and create new dance sequences. The techniques used in this process can recognize patterns in existing choreography and apply the found knowledge to create unique dance sequences that reflect a particular style. As examples, it

could include break dancing, ballet, or hip-hop. This process could revolutionize the field of dance composition, allowing dancers to create expressive and original dances much faster than traditional methods.

Aberman et al. [13] have discussed Deep video-based performance cloning as a technique used to capture and reproduce human performance, such as facial expressions, gestures, and movements. It uses deep learning algorithms to synthesize a clone of a person through videos or images. This technique can recreate an individual's entire performance from a video, replicating their exact facial expression, gestures, and movements. Deep video-based performance cloning has many applications, from virtual coaching to video game animation.

Zheng et al. [14] have discussed Time series data prediction and feature analysis of sports dance movements based on machine learning is a technology that can improve the accuracy, performance, and efficiency of predicting movements. It uses data that has been collected over some time in order to detect patterns and make predictions about future events. It can be used to predict everything from the action of an athlete when performing a specific move to the team's motion during a break in play. Using machine learning techniques can give coaches, players, and fans a better understanding of a sport's dynamics and help improve athletes' performance. It can also provide valuable insights into the pattern of movements associated with certain sports. By analyzing time series data, machine learning can also identify any discrepancies in an athlete's motion, such as variations in technique or forces. This technology can help coaches develop better training methods for their athletes and give players and fans a better understanding of the game they are playing.

Dong et al. [15] have discussed to improve the interpretability of deep neural networks with semantic information as a process of utilizing semantic information to enhance the behavior of deep neural networks. Deep neural networks are incredibly complex models, and understanding how they decide is difficult. Semantic information provides a way to make the decision-making process of a deep neural network more interpretable. Semantic information can add context to a model, helping explain why a specific decision was made. Semantic information can also reduce the complexity of explanations by categorizing results into more meaningful categories. This increased interpretability of deep neural networks ultimately helps to improve their performance and trustworthiness.

Sünderhauf et al. [16] have discussed the potential of deep learning for robotics is limitless. It can help in motion planning, visual perception, navigation, and decision-making. Deep learning algorithms can assist in advanced obstacle avoidance, automated driving, and home automation. The limits to deep learning in robotics are cost, infrastructure, and access to data. Cost is a primary limiting factor since deep learning algorithms require a lot of computing power to work, which can drive up the cost. Additionally, deep learning requires data to train appropriately, which can take time to acquire, depending on the industry. Finally, access to training and infrastructure can be a limitation as it can require specialized personnel with the right expertise.

Wu et al. [17] have discussed Image Comes Dancing With Collaborative Parsing-Flow Video Synthesis is a technique for creating videos from a single image. The technique involves using computer-based synthesis algorithms to analyze the source image, interpret it according to specific parameters, and produce a resulting video based on the input image. This technique has been used to create videos of dancing characters from a single image, with the characters dynamically dancing to match the rhythm of the music. It has also been used in various interactive video games and video-editing applications.

Kico et al. [18] have discussed Digitization and visualization of folk dances in cultural heritage is the process of digitally capturing, preserving and sharing the stories and movements of a variety of folk dances around the world. It can be achieved by recording and documenting folk performances, utilizing digital media, and creating interactive 3D computer visuals and animations. Digitization and visualization also help preserve and share the cultural heritage of folk dances for future generations, allowing for new interpretations of traditional motions. Ultimately, it serves as a way to spread cultural appreciation and create a bridge between different generations.

S. T. Ahmed et al. [19] have discussed dance style transfer with a cross-modal transformer is an artificial intelligence that takes insights from a source of dance-related material, such as online videos, and translates them into a different dance style. It does this by leveraging machine learning and natural language processing to try and capture the "essence" of a particular dance style. This technology has allowed for the rise of creative AI projects that can translate music, movie dances, or even user-created dance videos into different dance styles.

Cai et al. [20] have discussed An Automatic Music-Driven Folk Dance movement generation Method Based on a Sequence-To-Sequence Network is an AI-based approach to automatically generate folk dance movement

sequences by predicting and following musical input. This method uses a deep learning model incorporating music and motion data to generate more realistic and creative folk dance movements according to the musical beat. This method utilizes a sequence-to-sequence network architecture, which allows for creating new dance sequences by using various input sources. This method is beneficial for quickly creating and customizing folk dance performances with different musical genres. It can also provide a more reliable and entertaining way to generate custom dances than traditional choreographers.

Li et al. [21] have discussed Human motion recognition in dance video images based on attitude estimation is a technique used in computer vision and visual recognition to identify the body posture of a dancer in a video image. This technique uses a feature-based attitudinal translation architecture to detect and classify the body poses of a dancer from a given dance video sequence. The primary goal of human motion recognition in dance video images based on attitude estimation is to accurately recognize the body poses of a dancer in an unstructured dance environment to facilitate automated motion analysis and evaluation of dance performances.

**2.1. Research Gap.** In this model, the number of low power and low cost sensor has been deployment as self-adaptive and traffic dependent network protocol on the traffic of the network. The node data transmissions of the sensed data are adaptively changes to the traffic pattern. Power changes occur based on traffic [9]. The node will be time- synchronized for path negotiation and data contention on basis of node density. Path Allocation model of the protocol enhances the transmission capabilities on the less utilized nodes to prevent network degradation. Further linear programming architecture has been employed to Dynamic node hopping sequence. The routing architecture provides optimal stability among the node transmission time with respect to node availability and energy consumed on the effective path within specified delay along throughput constraints to solve energy whole problem

- Limited research on combining motion capture with deep learning techniques for generating the motion of characters derived from dance actions and more work must be done on implementing choreographic structures within coherent action generation.
- Few studies have examined the use of motion capture and behavior recognition techniques to monitor and evaluate the effectiveness of a generated action sequence in a human-robot interaction.
- Few studies have explored the effects of different music styles on dance-style action generation and little work needs to be done on adapting generated motion to context and environmental factors.
- The use of AI to enable autonomous, cooperative, and emergent motion patterns is still in its infancy and research is needed to develop machine learning models that capture temporal dynamics better and create realistic and natural-looking motion.

## 2.2. Research Objectives.

- To explore and generate novel dance movements and sequences through advanced deep learning architectures and strategies for dance motions.
- To develop an AI-based model, that is capable of creating movement variations to generate unique and appealing dance choreography
- To investigate the affective context of the dance motion by training different deep learning networks and gains an understanding of the underlying principles of motion in human-like dancing styles.
- To discover the different factors that influence dancing/motion in different rhythms and beats and design appropriate strategies for automated dance generation with respect to small details of the movements.
- To explore methods to improve the fluency and accuracy of predicted dance motions and construct an effective and efficient training environment for deep learning models to generate dancing sequences.

**3. Methodology.** The proposed dance coherent action generation model based on deep learning is a model for the automated generation of dance moves from videos. It applies deep learning to extract motion features from videos that represent the movement of dancers and use these features to generate new dance moves. The dance coherent action generation system has shown in the following fig. 3.1. The model is designed to generate smooth transitions between different moves and generate new moves that align with the video's overall style. It can also control the speed of the generated dance moves, creating high-quality visualizations of dancing videos. The model is trained on various dance data sets, which helps provide more accurate results.

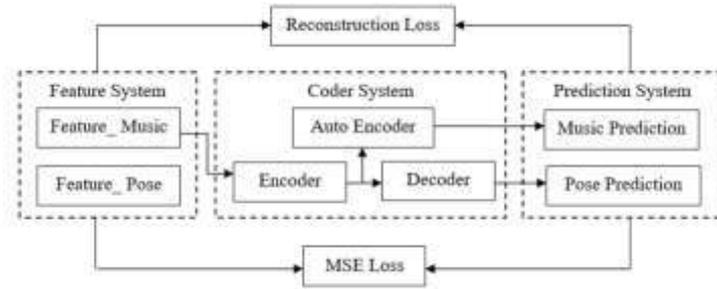


Fig. 3.1: Dance coherent action generation system

**3.1. Feature system.** It is a dance coherent action generation model extracts meaningful features from the environment. These features are used to inform the model of the state of the environment and to guide the decision-making process for artificial agents. These features can range from visual cues such as colors, shapes, and textures to auditory cues such as music tempo and pitch and kinetic features such as movements. The system also allows feature extraction from multiple domains, such as vision, audio, and motion. It allows for more robust and adaptive decision-making for the artificial agents in the model. Furthermore, feature extraction can detect global or local environmental patterns, making it easier for the model to generate dynamic and fluently choreographed dance sequences.

**3.2. Coder system.** It is an essential to the dance coherent action generation model. It performs a vital role in generating all the actions that constitute a dance. Its primary function is to encode the motion and music information into a format the system can understand. It does this by extracting meaningful visual and audio data from the dance video and associating it with different labels representing the motions and notes. The coder system then sends this encoded data to the agent, which then uses it to generate the sequences of actions that will constitute the dance. It ensures that the generated action will be coherent and consistent without discrepancies. The coder system also allows the model to “tune itself” by adjusting the resolution of the actions in order to reflect the desired motion better.

**3.3. Prediction systems.** It has dance-coherent action generation models are responsible for predicting the upcoming action of a dancer in order to generate a coherent dance routine. These systems are used to generate an output, which is a sequence of movements that correspond to a specific dance style, and it aims to produce a more realistic performance compared to pre-programmed dance steps. Prediction systems are used to interpret and anticipate the body movements of a dancer and generate appropriate sequences of movements in response. They use deep learning, statistical analysis, and motion capture techniques to track the dancer’s movement. Furthermore, they use the previous movements to anticipate the next movement, making it easier for the model to generate a coherent and connected dance routine for the dancer.

**3.4. Reconstruction loss.** It is a function used in the Dance Coherent Action Generation Model. It is used to optimize the model according to the human motion data that is being used. Reconstruction loss works by computing the difference between the model’s output and human motion data used for the input. The model adjusts its outputs by weighting and reducing the error resulting from the difference between the expected output and the actual human motion data. It means the model is more accurate, allowing it to predict human motions better. Reconstruction loss allows the Dance Coherent Action Generation Model to predict human motion better, providing more accurate output for the user.

**3.5. MSE (Mean Squared Error) loss.** It is a widespread loss function frequently used in machine learning models. In a dance coherent action generation model, this loss function calculates the errors between the predicted result and the real action. It measures the average of the squares of the errors or deviations, which is then used to gauge the model’s efficiency and improve its accuracy over time. This loss function is commonly used in supervised learning tasks like the dance coherent action generation model. By taking the

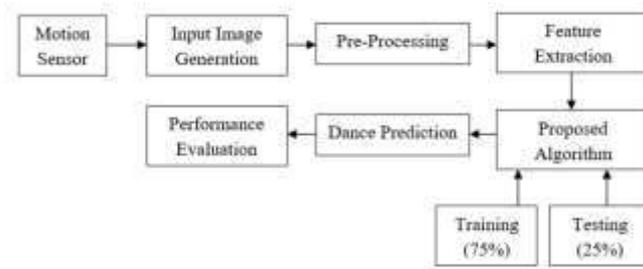


Fig. 4.1: Proposed block diagram

differences in prediction as errors, MSE loss can lead to improved predictions and more accurate generation of dance coherent actions.

A “coherent action” is a behavior that results in a composition of physical movements that is organized, stress-free, and capable of producing unified motion. It is a physical expression that is purposeful and organized, connecting body parts in an intrinsic manner that is relatively free from perceived tension. This definition specifically applies to the field of dance, where movement not only expresses emotion but also melds space, time, and music into a storytelling experience. In the field of dance, sensors, and sensor data allow information to be collected on how a dancer moves through space and can be used to create and execute more precise or customized actions for a choreographed routine. Sensors that can be used range from physical to digital and are used to create an interactive performance environment. The most commonly used physical sensors are motion-capture systems, accelerometers, inclinometers, and gyroscopic sensors. Digital sensors include cameras, microphones, and microphone arrays. These sensors and sensor data are used to generate a visualization of a dancer’s body, creating a visualization of the dancer’s movement, body shape, gesture, and interaction. This data is then analyzed and evaluated to produce an action that is meaningful and articulate. Ultimately, with the assistance of sensors and sensor data, a dancer’s routine can be designed, and the body position of the dancer can be monitored, resulting in a more precise action and a more accurate representation of the idea being expressed.

**4. Proposed Model.** The Dance Coherent Action Generation Model Using Deep Learning generates realistic, engaging, and fluid movement using Deep Neural Networks. It allows for a wide range of motion styles, from traditional hip-hop and freestyle movement to contemporary styles like jazz and Lyrical. The proposed block diagram has shown in the following fig. 4.1.

The model works by taking in raw video or motion-capture data of a dancer or performer’s body movements and producing motion-capture data that can be used for animation. It can fuse styles seamlessly, allowing for a much more comprehensive range of motion styles than traditional animation tools. The model can also divide a single movement style into multiple parts, allowing for speed and elegance. Finally, it can create new motion combinations from existing movement data to help innovate traditional dance styles. The proposed framework is shown in Table 4.1.

The proposed model can be applied to generate different types of dance styles by using specific data sets, such as motion capture data or motion synthesis, depending on the type of dance. This model can also be used for generating natural motion sequences, such as those seen in expressive dance. The proposed model should be evaluated against other existing models to analyze its efficacy in action generation. Existing action generation models can be used for comparison. It is also important to compare the qualities of the generated motion sequence with those of actual dance styles. The model should be evaluated in regards to its ability to be adapted to different body types and dancing abilities. Different body types and motion preferences can be encoded in the model to generate meaningful and personalized dance movements. Additionally, using different kinematic data from different dance skills can help refine the generated motion sequences to match desired levels of difficulty. Scalability is an important factor to consider when developing models. When expanding a dataset or model, one typically wants to do so in a manner that is as efficient as possible. The following are

Table 4.1: Dance Coherent Action Generation framework

1.	IP: a //Initial dance action;
2.	OP: b //Final dance action;
3.	Start
4.	GET_Sensor_IP ( ); //Get the sensor input data;
5.	Compute the a (t) ;
6.	if t=1
7.	PoP (t) = Compute the P ( ); // Computation of preliminary population;
8.	While (stop_con( ) ) do
9.	EVA_ the PoP (t); // Evaluation;
10.	CAP_IDA; // Capture the initial dance action;
11.	ADJ_F(a); // Adjust the fitness function;
12.	SEL_Next ( ); // Select the next action;
13.	GEN_PoP <sub>re</sub> ( ); //Generate repeat dance movements;
14.	GEN_PoP <sub>co</sub> ( ); //Generate cross over movements;
15.	GEN_PoP <sub>mc</sub> ( ); //Generate motion capture function;
16.	PoP (t+1) = (PoP <sub>re</sub> )+(PoP <sub>co</sub> )+(PoP <sub>mc</sub> );
17.	t = t+1; //Incremental operation;
18.	End;

some basic strategies that can be used to ensure scalability:

**Automated solution:** Automated solutions such as auto-scalers and autoencoders can be used to enable the model to scale up or down as the dataset grows or shrinks.

**Parallelization:** Parallelizing the dataset or model allows for the completion of tasks in much shorter timeframes while keeping scalability efforts to a minimum.

**Establish a baseline:** Establishing a baseline provides an organized structure upon which your model can be built. It also gives you a place to compare and identify areas where you can improve your scalability efforts.

**Preprocessing:** Preprocessing data can help simplify a dataset, making it easier to interpret and use for predictions, which can help the model become more scalable.

**Feature Selection:** Automated feature selection can help improve accuracy and reduce computational complexity while maintaining scalability.

**Model Evaluations:** Regular evaluations of the model to assess scalability can help determine what strategies need to be implemented (or improved) in order to achieve desired scalability.

**Observed sequences of actions:** Leveraging observed sequences of actions can provide predictive models that can be used to identify valuable insights from the data. It can allow for more informed decisions to be made and is an important tool for scalability.

The model is not necessarily robust enough to generate coherent actions in various environments with different sensor configurations. Different environments and sensor configurations might require different approaches to agent behavior. Further fine-tuning of the model would be necessary to make it robust enough for various environments.

**4.1. Pre-processing.** It is a significant step in proposed model and is especially beneficial for a dance-coherent action generation model. Pre-processing can help to normalize the input data, clean up noise, and reduce redundancies. Most importantly, it can reduce the dimensionality of the input data, making it easier for the algorithm to learn a profound representation of the data. The upper bound dance actions has indicated in the following eq. 4.1

$$\forall a > 0; \forall a \in X \Rightarrow x \leq \text{feature\_}X; \quad (4.1)$$

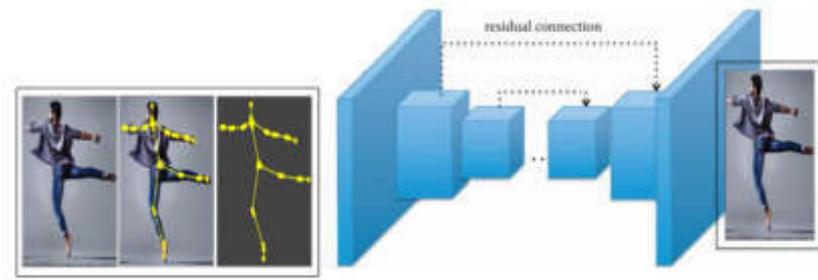


Fig. 4.2: Dance Image Generator

where  $a$  is an initial dance action,  $X$  is an upper bound dance segment and  $x$  is a coherent function. The lower bound dance actions has indicated in the following eq. 4.2

$$\forall b > 0; \forall b \in Y \Rightarrow y \leq feature\_Y; \tag{4.2}$$

where  $b$  is an final dance action,  $Y$  is a lower bound dance segment and  $y$  is a coherent function. The convexity function has shown in the following,

$$UB : q(\alpha s_1 + (1 - \alpha)s_2) \leq \alpha q(s_1) + (1 - \alpha)q(s_2) \tag{4.3}$$

$$LB : q(\alpha s_1 + (1 - \alpha)s_2) \geq \alpha q(s_1) + (1 - \alpha)q(s_2) \tag{4.4}$$

where UB shows the upper bound dance action and LB shows the lower bound dance action. The Dance Image Generation has shown in the following fig. 4.2.

Pre-processing can also help fill any gaps in the data or scale the features so that different data sources can be used together. Finally, it can reduce the computational complexity associated with complex deep neural networks, allowing the model to more quickly and effectively. The sigmoid function has shown in the following eq. 4.5

$$A(z) = \left\{ \frac{1}{1 + e^{-z}} \right\} \tag{4.5}$$

The image generator in the dance coherent action generation model using deep learning is a convolutional neural network-based architecture. This model uses a long short-term memory (LSTM) layer as a guide for generating image frames, which are then given as inputs to the convolutional neural network. The directional derivatives has shown in the following,

$$\frac{\partial q}{\partial d} = \frac{\partial q}{\partial s} \cos \phi + \frac{\partial q}{\partial c} \sin \phi \tag{4.6}$$

The convolutional neural network then passes the image frames through convolutional layers, which extract features. These features are then used to generate the next image frame. This process is repeated until the entirety of the motion is generated. The model then produces a sequence of motion vector representations used to reconstruct the motion as a video clip or animation. The linear interaction has shown in the following eq. 4.7

$$a_x = \beta(C_e * [d_{x-1}, f_x] + g_e) \tag{4.7}$$

$$b_y = \beta(C_e * [d_{y-1}, f_y] + g_e) \tag{4.8}$$

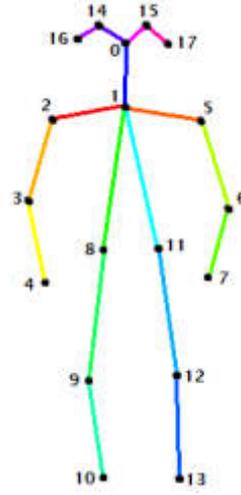


Fig. 4.3: Human body feature points

where  $a$  and  $b$  are the dance action,  $\beta$  is the coherence factor,  $x$  and  $y$  are the motion sensor coordinates.  $C$  is the coder function. The image generator in a deep learning model for dance coherent action generation visualizes the input movements. This visual representation is then used to generate a smooth, coherent sequence of output movements. It serves as an efficient way to efficiently produce an action from the input motion, thus preserving the continuity of the motion. Additionally, the image generator is used to create a visual representation of the output sequence used as a reference to compare the motion generated by the model.

**4.2. Feature Extraction.** Feature extraction in deep learning is extracting meaningful features from data. It is a crucial step in any deep learning model, as it helps to identify patterns in the data and determine which features are essential to capture. In proposed model, feature extraction is used to create valuable representations of motion sequences, which can then be used as input for the model.

$$a_x = \beta(C_e * d) \quad (4.9)$$

$$b_y = \beta(C'_e * d') \quad (4.10)$$

$$C' = C^T \quad (4.11)$$

$$L(ab) = \|a - b\|^2 \quad (4.12)$$

$$L_X(a, b) = - \sum_{m=1}^n \{a_m \log b_m + (1 - a_m) \log(1 - b_m)\} \quad (4.13)$$

where  $a$  is the initial dance action,  $b$  is the final dance action,  $C$  represent the coder function. It allows the model to identify better and generate the correct dance moves. By using feature extraction on motion sequences, the model can better differentiate between features important to replicating a dance move and those not. Fig. 4.3 shows the human body feature points.

The human body feature points in the dance coherent action generation model refer to points on the body detected by a camera and used to predict movements and generate performers' dance motions. These points are typically used to capture motion and are placed on essential body parts, such as the arms, legs, torso, and head joints. Table 4.2 shows the human body feature points in detail.

Table 4.2: Human body feature points

S.No	Feature Point	Representation
0	Dancer Nose	a1, a2
1	Dancer Neck	b1, b2
2	Dancer Shoulder (Right)	c1, c2
3	Dancer Elbow (Right)	d1, d2
4	Dancer Hand (Right)	e1, e2
5	Dancer Shoulder (Left)	f1, f2
6	Dancer Elbow (Left)	g1, g2
7	Dancer Hand (Left)	h1, h2
8	Dancer Leg (Right)	i1, i2
9	Dancer Knee (Right)	j1, j2
10	Dancer Foot (Right)	k1, k2
11	Dancer Leg (Left)	l1, l2
12	Dancer Knee (Left)	m1, m2
13	Dancer Foot (Left)	n1, n2
14	Dancer Eye (Right)	o1, o2
15	Dancer Eye (Left)	p1, p2
16	Dancer Ear (Right)	q1, q2
17	Dancer Ear (Left)	r1, r2

This information is then used to interpolate between poses and motions to generate a realistic and natural-looking output. For example, these feature points can represent a dancer's motion, such as twisting, turning, and stepping, and can generate a realistic and synchronized version of the dancer's motion. This way, the model can generate a natural and accurate representation of the dancer's motions. The improved training process in the Dance Coherent Action Generation Model using Deep Learning involves a two-stage approach. In the first stage, the model is trained on synthesized sequences of postures adapted from go-go dance videos. It allows the model to learn the general motion structure of the dance. In the second stage, the model is then fine-tuned using real-world data. This fine-tuning allows the model to model the nuances of real-world dancing and efficiently produce realistic results. The improved training process has shown in the following fig. 4.4.

As a result, the improved training process provides a more accurate representation of the physical body movements of go-go dancing. It ultimately helps in creating more realistic and coherent dance sequences.

**4.3. Detection.** Detection in dance coherent action generation models using deep learning allows the model to detect specific features in the input image or video sequence data used to generate dance steps autonomously. In the context of neural network models, detection typically refers to providing a model with the necessary visual cues required to generate the desired output. The detection of motions using the sensor has the following,

$$a_m = \sum_{x=1}^T \{\delta_{x,m} * g_x\} \quad (4.14)$$

Where,  $a$  is the input dance detection in  $m^{th}$  state,  $\delta$  is the attention weight,  $x$  and  $m$  are the inputs of deep learning model. These visual cues can be in the position, size, or orientation of particular objects in the image. In a dance setting, the model must detect the dancer's body position, movement patterns, and the context of the dance scene to generate coherent movements. To obtain the detection of dimensional dance moves has the following,

$$L^x = \mu \{C, L^{x-1}\} \quad (4.15)$$

$$R^{T_x} = \omega^x \{C, L^{T_x}\} \quad (4.16)$$

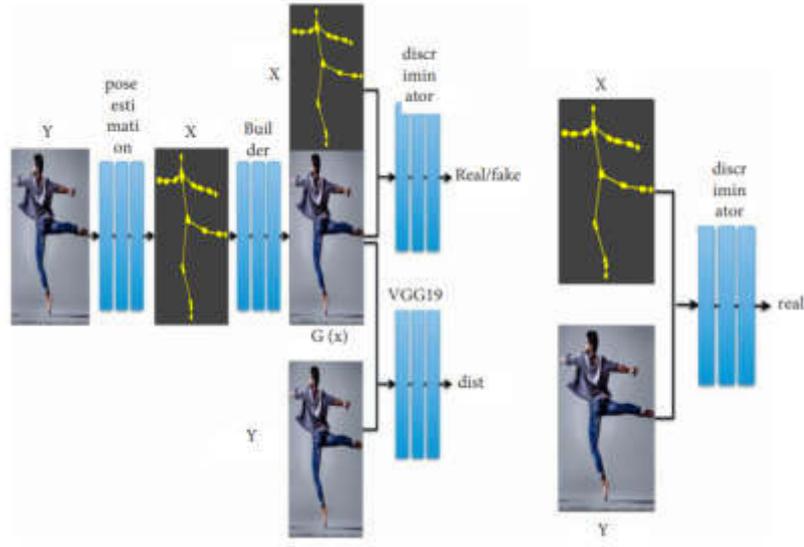


Fig. 4.4: Improved training process

$$R^T = \omega^x \{C, L^{T_x}, R^{x-1}\} \quad (4.17)$$

$$F = \min_a \max_b L_x(a, b) + \alpha_x L_{MSE}(a) + \alpha_y L_x(E, D) \quad (4.18)$$

where  $F$  is the final detection of dance coherent action,  $a$  is the initial dance action and  $b$  is the final dance action. MSE shows the mean square error,  $E$  is the encoder and  $D$  is the decoder. The proposed model may be used to detect the location and movement of the dancer's limbs, which can then be used to generate the best set of moves that synchronize with the music. Detection is necessary for deep learning models as it allows them to acquire visual cues to generate the desired output.

**4.4. Classification.** Classification in dancing is used to categorize the types of movements a dancer performs [22]. Regarding consistent action generation using deep learning, classification allows the model to identify and recognize patterns among dance moves and use that knowledge to generate new, similar moves. Dance Coherent Analysis is an important part of the Coherent Action Generation Model using Deep Learning. This model uses deep learning to generate realistic motion of a dancer that is both expressive and organic, while still being physically and aesthetically pleasing. The Dance Coherent Analysis module consists of an analytical pipeline to detect and analyze the motion data of the dancer. It extracts features related to body positions and movements from the footage, estimates the dancer's plans, and extracts motion metrics that can be used to describe the performance of the dancer. The extracted features are then used to generate a trained behavior model for a dancer. The resulting model is used to animate a dancer in a simulated world, to generate dance-based applications. This model helps animate realistic dancing motions with few parameters and makes it much easier to generate realistic animations. From eq.18, the  $L_x(a,b)$  and  $LMSE(a)$  has classified as the following,

$$L_x(a, b) = F_{(c,d)} \{\log b(c, d)\} + F_c \{\log(1 - b(c, a(c)))\} \quad (4.19)$$

$$L_{MSE}(a) = F_{(c,d)} \{\log b(c_{e-1}, c_e, d_{e-1}, d_e)\} + F_c \{\log(1 - b(c_{e-1}, c_e, a(c_{e-1}), a(c_e)))\} \quad (4.20)$$

The dance coherent analysis has shown in the following fig. 4.5.

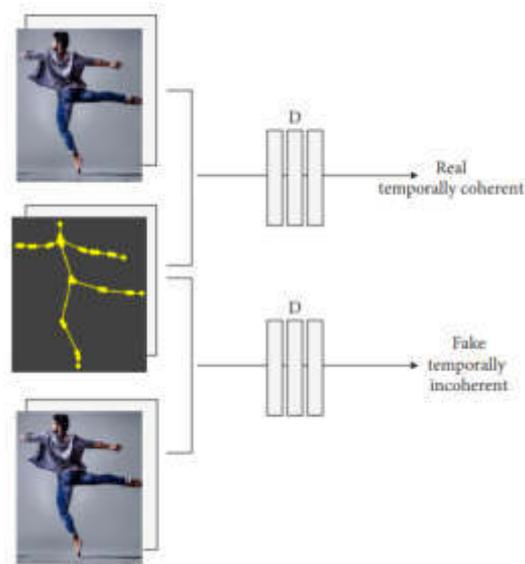


Fig. 4.5: Dance coherent analysis

Classification also helps the model recognize and understand the elements of the dance moves to generate new steps that fit within the constraints of the choreography. It allows the model to produce smooth, realistic dancing sequences that feel natural and coherent.

Sensor-based dance coherent action generation model using a deep learning framework is an innovative approach to using a combination of sensing, processing, and predictive algorithms to generate ground-breaking, impactful, and expressive human-like dances. By combining sensing and deep learning techniques, this model allows for optimized and automated processes of sensor data, which can enable truly autonomous generated dances. The model employs algorithms to learn the trends in motion performance from dance sequences, the variations in movements, and the correlations between different movements. It provides a potential performance outcome on the basis of the captured motion data in the form of a probability distribution instead of a single prediction. This approach improves the motion performance, fosters an emotionally expressive performance, and increases efficiency while reducing user input. This model is also scalable to multi-sensor dance training, making it a powerful tool for creating larger-scale, intelligent, and expressive dances through machine learning.

Learning systems that are able to detect relevant information from sensors and generate coherent actions accurately are typically based on artificial neural network algorithms. These systems use the input from the sensors to learn the characteristics associated with a particular type of behavior. Once the system has enough data, it can be trained to recognize similar patterns and generate actions accordingly. For example, an object recognition system can identify different objects in an image and process this information so that it can take an appropriate action. Additionally, reinforcement learning algorithms can be used to teach an agent to find the optimal path to a given goal based on various types of feedback from the environment. With the help of supervised learning algorithms, the system can learn to recognize specific patterns and generate appropriate responses.

**5. Comparative Analysis.** The proposed dance coherent action generation model (DCAGM) has compared with the existing Recurrent Neural Network (RNN), Dance Based on Deep Learning (DBDL), Quantum-based creative generation method (GCGM) and Convolutional sequence generation (CSG). Here python is a simulation tool used to execute the results.

**5.1. Computation of Accuracy.** Accuracy refers to the percentage of correctly classified examples in a deep learning dance coherent action generation model. The number of correctly classified examples must be

Table 5.1: Comparison of Accuracy (in %)

No.of Inputs	RNN	DBDL	GCGM	CSG	DCAGM
100	65.05	58.72	72.07	74.22	97.46
200	63.56	56.75	69.65	72.02	95.47
300	62.76	55.62	69.24	71.22	94.27
400	60.43	54.41	67.64	70.55	93.79
500	59.42	54.04	65.32	69.12	92.36
600	58.78	52.51	64.07	68.03	92.20
700	58.12	52.01	61.34	67.55	91.43

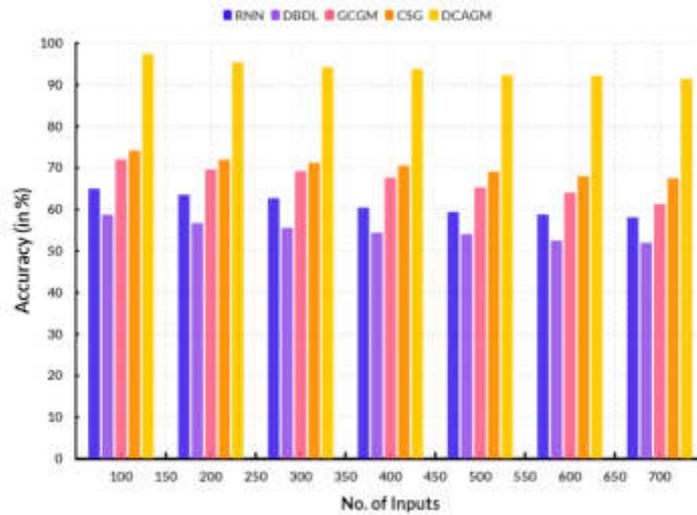


Fig. 5.1: Accuracy

divided by the number of examples examined. The comparison of accuracy has shown in the following table 5.1. It is a widespread loss function frequently used in machine learning models. In a dance coherent action generation model, this loss function calculates the errors between the predicted result and the real action. It measures the average of the squares of the errors or deviations, which is then used to gauge the model's efficiency and improve its accuracy over time. This loss function is commonly used in supervised learning tasks like the dance coherent action generation model. By taking the differences in prediction as errors, MSE loss can lead to improved predictions and more accurate generation of dance coherent actions.

Fig. 5.1 shows the comparison of accuracy. In a computation tip, existing RNN obtained 60.43%, DBDL obtained 54.41%, GCGM obtained 67.64%, and CSG obtained 70.55% accuracy. The proposed DCAGM obtained 93.79% accuracy. The proposed model does make use of transfer learning techniques to improve the accuracy of action generation. Transfer learning involves utilizing information and knowledge gained from existing models to improve the performance of new models. Transfer learning can be used to take data from previously trained models and apply it to new tasks, helping the model learn without having to start from scratch. In the proposed model, the transfer learning technique is used to help the model learn more quickly and accurately from previous data to improve its performance when generating new actions.

**5.2. Computation of Precision.** The computation of precision has to predict the action sequences and the actual test data used to train the model. This task is usually done by calculating the overall accuracy, which is the proportion of correctly identified correct actions versus the total number of observations in the test

Table 5.2: Comparison of precision (in %)

No.of Inputs	RNN	DBDL	GCGM	CSG	DCAGM
100	75.05	68.72	82.07	84.22	97.46
200	73.56	66.75	79.65	82.02	95.47
300	72.76	65.62	79.24	81.22	94.27
400	70.43	64.41	77.64	80.55	93.79
500	69.42	64.04	75.32	79.12	92.36
600	68.78	62.51	74.07	78.03	92.20
700	68.12	62.01	71.34	77.55	90.43

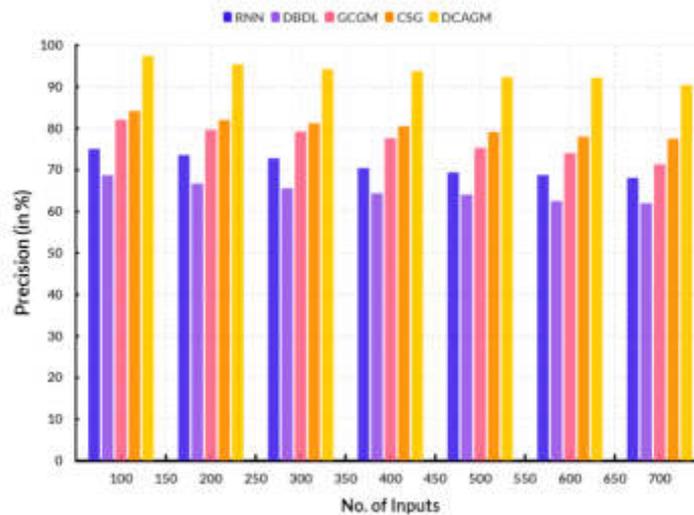


Fig. 5.2: Precision

data. In addition to the overall accuracy, precision of the predictions is also measured. Precision reflects the ability of the model to identify the correct dance actions with positive examples only correctly. The comparison of precision has shown in the following table 5.2.

Fig. 5.2 shows the comparison of precision. In a computation tip, existing RNN obtained 70.43%, DBDL obtained 64.41%, GCGM obtained 77.64%, and CSG obtained 80.55% precision. The proposed DCAGM obtained 93.79% precision.

**5.3. Computation of Recall.** The recall of a model in dance coherent action generation using deep learning is calculated by taking the ratio of the total number of correctly predicted classes to the total number of authentic classes in the dataset. It measures how many of the actual classes were predicted correctly by the model. A higher recall value indicates that the model is more likely to classify unseen data correctly. The comparison of recall has shown in the following table 5.3.

Fig. 5.3 shows the comparison of recall. In a computation tip, existing RNN obtained 62.26%, DBDL obtained 57.63%, GCGM obtained 79.42%, and CSG obtained 75.33% recall. The proposed DCAGM obtained 97.75% recall.

**5.4. Computation of F1-Score.** The F1-score for a Dance Coherent Action Generation Model using deep learning measures how accurately the model predicts the dancer's movement or how well the model has been trained. It is composed precision and recall. Precision measures the accuracy of the predictions made

Table 5.3: Comparison of recall (in %)

No.of Inputs	RNN	DBDL	GCGM	CSG	DCAGM
100	61.01	57.53	77.55	75.64	97.53
200	61.51	57.53	78.64	75.90	97.64
300	62.26	58.36	79.78	76.47	97.70
400	62.26	57.63	79.42	75.33	97.75
500	61.21	56.52	77.89	74.31	97.79
600	60.93	56.12	77.25	74.07	97.82
700	61.65	56.69	77.83	74.72	97.84

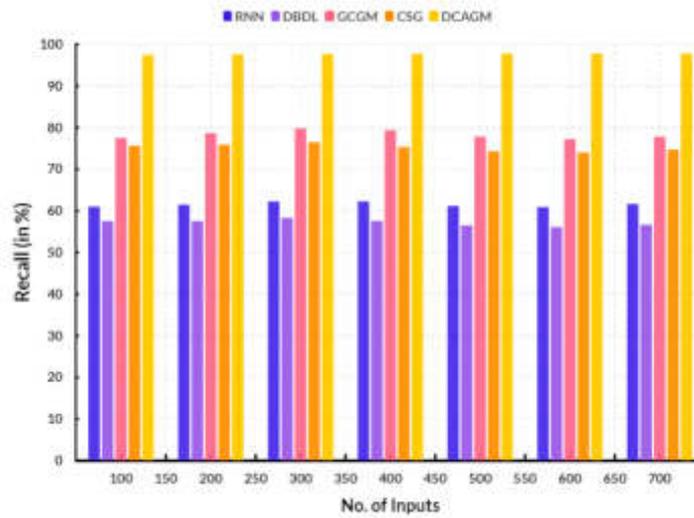


Fig. 5.3: Recall

by the model. The recall measures how well the model can identify all of the movements that the dancer is performing. By combining these two metrics, the F1 score indicates how accurately the model predicts the dancer's actual movement. The comparison of F1 score has shown in the following table 5.4

Fig. 5.4 shows the comparison of F1 score. In a computation tip, existing RNN obtained 72.46%, DBDL obtained 68.26%, GCGM obtained 74.46%, and CSG obtained 83.36% F1 score. The proposed DCAGM obtained 92.92% F1 score. The proposed dance coherent action generation model using deep learning has several advantages over traditional methods of choreography and animation:

- The model uses a deep learning approach to learn the concept of dancing from existing examples. It enables the model to capture the subtle nuances of a particular dance style, such as musicality, emotion, and the human-like movements that come with the style.
- The model can generate high-quality motions that can be repeated later with little or no modifications. It means that the motions generated by the model are less likely to drift from their representations over time, making them more suitable for use in live performances.
- The model also has the potential to adapt to changes in the environment or the music to which it is dancing, enabling the user to make minor adjustments to the dance motion generated without needing to start from scratch.

The proposed dance coherent action generation model using deep learning is limited in two main ways:

- The current deep learning models are primarily supervised and require large datasets of labeled exam-

Table 5.4: Comparison of F1 score (in %)

No.of Inputs	RNN	DBDL	GCGM	CSG	DCAGM
100	66.85	64.83	67.96	79.01	88.45
200	68.52	65.96	70.89	80.27	90.92
300	70.47	66.31	72.43	82.16	91.72
400	72.46	68.26	74.46	83.36	92.92
500	75.04	69.03	75.36	84.92	93.56
600	77.03	69.41	77.33	86.67	94.82
700	79.05	70.54	78.80	87.60	95.82

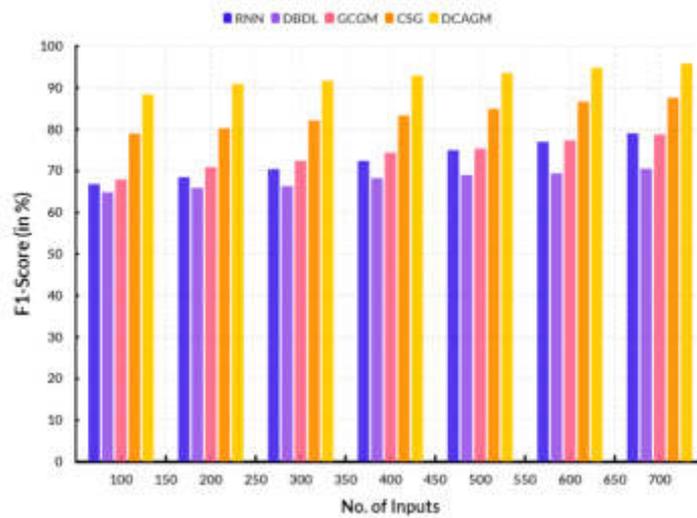


Fig. 5.4: F1 score

ples, which may be expensive and difficult to obtain.

- The models are often over-sensitive to noisy input data, making them unable to generalize and prone to over fitting.

The coherent method of the Sensor-Based Dance Coherent Action Generation Model using the Deep Learning Framework improves the existing models in several ways. Firstly, the coherent method makes use of sensors with different modalities (audio, visual, etc.) to capture the 3D coordinates (x,y,z) of each body part and to calculate the relative distances between body parts of the dancer. The data thus captured is then used to build a 3D action space model that captures motion data more efficiently and compactly compared to 2D models. This 3D action space model also improves the accuracy of the generated action sequences as it is able to capture more complex motions with greater precision. In addition, the use of deep learning algorithms to generate the action sequences helps to provide more accurate predictions while having reduced training time. Finally, the use of a real-time evaluation system helps to ensure that the generated action sequences are correct and that no mistakes are made during the generation process. All of these benefits together help to improve the existing sensor dance coherent action generation model using deep learning framework, making it more efficient, accurate, and reliable.

The proposed framework can be extended to multiple dance styles by using additional motion capture data from different dance style motions. This extended dataset could be used to train the network better and create a more generalizable model for different styles of dance. It could provide more accurate results when transferring

motions from one style of dance to another. The results from this proposed framework have shown that it is possible to transfer motions between two different style movements and maintain the stylistic characteristics of the original motion. It has been achieved by preserving the continuity of the motions while modifying the amplitude and timings to fit the new style. Future research could explore various ways to improve the accuracy of this framework, including using larger datasets or adding noise to the motions to make the framework more robust. Further research could investigate the possibilities of using visuals (e.g., avatars) in addition to motion capture data to transfer both the motion and visual style of different dance styles. It could enable the framework to provide more realistic motion transfers between different dance styles.

**6. Conclusions.** The dance coherent action generation using deep learning is a feasible and promising approach to generating natural, realistic, and choreographically composed actions. Deep learning has enabled the autonomous generation of dance motions that are inherently choreographed and expressive. It has generated high-quality and expressive actions for various dance styles, including modern, tap, street dance, ballroom, and Latin. In addition, the generated motions demonstrate great diversity and variation, allowing for the implementation of a procedurally generated scenario in which motion is generated from a given scenario. The proposed model obtained 93.79% accuracy, 93.79% precision, 97.75% recall and 92.92% F1 score. The future scope of dance coherent action generation using deep learning is comprehensive due to the virtually limitless possibilities of combining Artificial Intelligence (AI) techniques such as deep learning with computer vision and natural language processing to create powerful machines capable of recognizing and generating complex choreography. It could be applied in various contexts, such as virtual tours, robotic dance groups, virtual reality performances, or even healthcare contexts for physical therapy and rehabilitation. With advancements in technology and imagination, the possibilities are endless. Sensor-based dance coherent action generation models using deep learning frameworks can be used in applications such as automatic motion generation and analysis. It can be applied to various robotic applications such as character animation for animated movies and games or even for industrial robotics for precise and efficient operations. For future research, this model can be used to create more complex, sophisticated, and realistic motion sequences. It can be used to develop motion generation methods that can learn from complicated interactions between different actors in the environment and for detecting and recognizing motions in a natural human motion. It can also be used to extend and improve the performance of existing motion recognition and generation methods. This model can be used to explore using motion to further improve current deep learning frameworks, especially for tasks like image segmentation and classification.

#### REFERENCES

- [1] Ma, X., & Wang, K. "Dance Action Generation Model Based on Recurrent Neural Network," *Mathematical Problems in Engineering*, 2022.
- [2] Kong, Y., & Fu, Y. "Human action recognition and prediction: A survey," *International Journal of Computer Vision*, vol. 130, pp. 1366-1401, 2022.
- [3] Ushiyama, J., Takahashi, Y., & Ushiba, J. "Muscle dependency of corticomuscular coherence in upper and lower limb muscles and training-related alterations in ballet dancers and weightlifters," *Journal of applied physiology*, vol. 109, pp. 1086-1095, 2010.
- [4] Lu, R. "Analysis of Main Movement Characteristics of Hip Hop Dance Based on Deep Learning of Dance Movements," *Computational Intelligence and Neuroscience*, 2022.
- [5] Maithili, K., Vinothkumar, V., & Latha, P. (2018). Analyzing the security mechanisms to prevent unauthorized access in cloud and network security. *Journal of Computational and Theoretical Nanoscience*, 15(6-7), 2059-2063.
- [6] Mei, P., Ding, G., Jin, Q., Zhang, F., & Jiao, Y. "Quantum-based creative generation method for a dancing robot," *Frontiers in Neurobotics*, vol. 14, pp. 559366, 2020.
- [7] Yan, S., Li, Z., Xiong, Y., Yan, H., & Lin, D. "Convolutional sequence generation for skeleton-based action synthesis," In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 4394-4402, 2019.
- [8] Ye, Z., Wu, H., Jia, J., Bu, Y., Chen, W., Meng, F., & Wang, Y. "Choreonet: Towards music to dance synthesis with choreographic action unit," In *Proceedings of the 28th ACM International Conference on Multimedia*, pp. 744-752, Oct 2020.
- [9] Hu, H., Liu, C., Chen, Y., Jiang, A., Lei, Z., & Wang, M. "Multi-Scale Cascaded Generator for Music-driven Dance Synthesis," In *2022 International Joint Conference on Neural Networks (IJCNN)*. pp. 1-7, July 2022.
- [10] Zhang, S., Yao, L., Sun, A., & Tay, Y. "Deep learning based recommender system: A survey and new perspectives," *ACM computing surveys (CSUR)*, vol. 52, pp. 1-38, 2019.
- [11] Muthukumar, V., Kumar, V. V., Joseph, R. B., Munirathanam, M., & Jeyakumar, B. (2021). Improving network security

- based on trust-aware routing protocols using long short-term memory-queuing segment-routing algorithms. *International Journal of Information Technology Project Management (IJITPM)*, 12(4), 47-60.
- [12] Zhang, X., Yang, S., Xu, Y., Zhang, W., & Gao, L. "Mining and applying composition knowledge of dance moves for style-concentrated dance generation," In *Proceedings of the AAAI Conference on Artificial Intelligence* (vol. 37, 5411-5419, June 2023).
  - [13] Aberman, K., Shi, M., Liao, J., Lischinski, D., Chen, B., & Cohen-Or, D. "Deep video-based performance cloning," In *Computer Graphics Forum*, vol. 38, pp. 219-233, May 2019.
  - [14] Zheng, D., & Yuan, Y. "Time Series Data Prediction and Feature Analysis of Sports Dance Movements Based on Machine Learning," *Computational Intelligence and Neuroscience*, 2022.
  - [15] Dong, Y., Su, H., Zhu, J., & Zhang, B. "Improving interpretability of deep neural networks with semantic information," In *Proceedings of the IEEE conference on computer vision and pattern recognition* , pp. 4306-4314, 2017.
  - [16] Sünderhauf, N., Brock, O., Scheirer, W., Hadsell, R., Fox, D., Leitner, J., ... & Corke, P. "The limits and potentials of deep learning for robotics," *The International journal of robotics research*, vol. 37(4-5), pp. 405-420, 2018.
  - [17] Wu, B., Xie, Z., Liang, X., Xiao, Y., Dong, H., & Lin, L. "Image Comes Dancing With Collaborative Parsing-Flow Video Synthesis," *IEEE Transactions on Image Processing*, vol.30, pp. 9259-9269, 2021.
  - [18] Kico, I., Grammalidis, N., Christidis, Y., & Liarokapis, F. "Digitization and visualization of folk dances in cultural heritage: a review," *Inventions*, vol. 3, pp. 72, 2018.
  - [19] Yin, W., Yin, H., Baraka, K., Kragic, D., & Björkman, M. "Dance style transfer with cross-modal transformer," In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pp. 5058-5067, 2023.
  - [20] Ahmad, M., Beddu, S., binti Itam, Z., & Alanimi, F. B. I. "State-of-the-art compendium of macro and micro energies. *Advances in Science and Technology Research Journal*. vol. 13, pp 88–109 <https://doi.org/10.12913/22998624/103425>, 2019.
  - [21] Cai, X., Xi, M., Jia, S., Xu, X., Wu, Y., & Sun, H. "An Automatic Music-Driven Folk Dance Movements Generation Method Based on Sequence-To-Sequence Network," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 37, pp. 2358003, 2023.
  - [22] Li, N., & Boers, S. "Human Motion Recognition in Dance Video Images Based on Attitude Estimation," *Wireless Communications and Mobile Computing*, 2023.

*Edited by:* Polinpapilinho Katina

*Special issue on:* Scalable Dew Computing for Future Generation IoT systems

*Received:* Sep 29, 2023

*Accepted:* Dec 13, 2023



## IOT BASED DANCE MOVEMENT RECOGNITION MODEL BASED ON DEEP LEARNING FRAMEWORK

ZHEN JI\* AND YAONONG TIAN†

**Abstract.** Deep Learning is becoming an emerging field in the Internet of Things (IoT) due to its ability to provide a comprehensive approach to automatic feature extraction and predictive modeling for analysis and decision-making. This paper introduces an IoT-based Dance Movement Recognition Model based on a Deep Learning Framework. The framework consists of a convolutional neural network (CNN) with a data-centric architecture to identify dance movements from the acquired data gathered by an IoT device. The IoT device collects 3D motion data captured by three accelerometers. Feature extraction is then done with the CNN architecture, resulting in a flattened matrix representing the movement. Subsequently, a Multi-Layer Perception (MLP) is used to classify the movements. The proposed system is experimentally evaluated on a standardized dataset of 16 dance steps with three-speed levels. The results show that our model outperforms state-of-the-art approaches in accuracy, evaluation time, and classification accuracy. The proposed model reached 90.74% accuracy, 87.12% precision, 83.78% recall and 84.39% F1-Score. The proposed model can serve as a basis for a reliable and intuitive system that can be used to monitor patient's dance movements with accuracy.

**Key words:** Deep learning, IoT, predicting model, dance, accuracy, precision, recall, f1-score

**1. Introduction.** IoT-based dance movement recognition is a system that uses Internet of Things technology to recognize dance movements when a person is dancing. It uses sensors to detect motion and sends the data to an analysis engine that can recognize different moves and generate reports of the person's performance [1]. This system can be used for educating dancers, assessing a professional dancer's performance, and even for games. It can also give dancers real-time feedback, helping them improve their technique. A residual connection is a type of shortcut in a neural network to allow the gradients in a network to flow more efficiently [2]. The residual connection in image generation helps bridge the gap between the generator and the discriminator by allowing the generator to utilize information from the discriminator to generate better images. The residual connections enable the generator to use the discriminator's features to generate higher quality and more coherent outputs [3]. It helps generate more semantically correct and meaningful output images that are appropriately categorized. It helps the discriminator generate accurate classifications. Allowing the generator to use information from the discriminator enables better dance coherent action generation [4]. Dance pose estimation in coherent action generation models using deep learning is the process of predicting dancers' movements from detected body poses. It is accomplished by combining deep learning algorithms such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs) [5]. By learning from sequences of previously recorded motion data, a deep learning model can detect various poses in dance actions and then generate dance sequences that are similar to them. It can significantly improve the realism of the generated actions and make them resemble those of authentic dancers. Furthermore, such models can be used in other applications like interactive gaming or robotic control [6]. The Dance Pose Builder takes a set of human-defined poses as input and then learns an efficient sequence of encoded representations by learning from the input data. It allows the generator to generate multiple sequences of meaningful movements, merging the poses in a controllable and creative manner [7]. The Dance Pose Builder allows for various motion descriptions to be inputted into the model and can be used to generate more complex dance movements that are still easily readable to a human audience. The Dance Pose Discriminator is a component of the Coherent Action Generation model using Deep Learning that determines whether a series of dance motion poses is generated by a trained model or by a natural dancer [8]. This function is critical in the training process, as it ensures that the model produces only realistic

---

\*International College, Krirk University, Bangkok, 10220, Thailand (corresponding author, [jizhen198321@sina.com](mailto:jizhen198321@sina.com))

†International College, Krirk University, Bangkok, 10220, Thailand ([tyn1075@hznu.edu.cn](mailto:tyn1075@hznu.edu.cn))

dance poses. The Discriminator also assists the model in learning how to transition from one pose to the next appropriately and realistically. It also helps to ensure that the correct sequence of poses is generated to create a coherent dance performance [9].

The Real Temporally Coherent (RTC) model is a deep learning-based action generation model for creating coherent dance movements. The model uses an Encoder-Decoder-like architecture, similar to the architecture used in natural language processing [10]. The encoder consists of a series of convolutional layers that take in the input motion capture data and generate an embedding vector that people the characteristics of the dance movement. The decoder then uses this vector to generate coherent and believable frames of animation of the dance movement outputted. The RTC model takes advantage of temporal information, allowing it to better recognize and reproduce the cyclic features often in dance movements, leading to more coherent and life-like output [11]. The Dance Coherent Action Generation model using Deep Learning involves generating temporally coherent dances from a single motion capture sample. It is accomplished using a Variation Auto encoder (VAE) for representation learning. The VAE maps the motion capture data into a fixed-sized representation. Then, a generative style transfer network transforms the extracted feature vector from the VAE into a temporally coherent dance sequence. This generative style transfer network is a convolutional neural network (CNN) trained on various motion capture data sets. The CNN/VAE setup is trained jointly in an end-to-end fashion, allowing the model to learn how to generate a temporally coherent output that is stylistically similar to the original motion capture sample [12]. After training the network, it can generate temporally coherent dancing sequences from any motion capture sample. The main contribution of the research has the following:

1. Improved accuracy of dance moves: With the Internet of Things (IoT), dance moves can be tracked more precisely than ever. It increases the accuracy of the recognition of specific dance moves.
2. Enhanced interactivity: By providing real-time feedback, the dancer can adjust their movements to get the best performance. It increases the overall interactivity of the dance experience.
3. Improved safety: IoT-based recognition systems can notify the dancer if they are performing a move incorrectly, which is beneficial to both the dancer's safety and preventing injuries.
4. Enhanced customizability: IoT-enabled devices allow dancers to customize their dance moves or create entirely new ones, adding creativity and customizability to the dance experience.
5. Wider applications: IoT-enabled devices can be used in various activities, including exercise tracking, rehabilitation, video games, and many more. It increases the number of ways in which the technology can be applied.
6. Remote monitoring: IoT-enabled systems can allow dancers to remotely monitor their progress in real-time or by permanently recording all their dance moves. It allows for constant feedback and progress tracking

**2. Related Work.** Zhang et al. [13] has discussed an Optimization simulation of the match between technical actions and music of a national dance based on deep learning is an advanced artificial intelligence technology that uses deep learning algorithms to create an algorithmic system that can accurately simulate the matching of technical actions and music of a national dance. The simulation is optimized using statistics, signal processing, and computer vision techniques. This technology can help choreographers train their national dance teams and evaluate and compare the performances of different dancers.

Karthick Raghunath, K. M et al. [14] has discussed the Time Series Data Prediction and Feature Analysis of Sports Dance Movements Based on Machine Learning uses artificial intelligence, such as neural networks and support vector machines, to identify patterns in dance movements. These patterns can be used to accurately forecast future sports dance movements and determine correlations between various movements. Feature analysis involves extracting important features from recorded sports dance movements, such as motion, speed, and acceleration. With the help of machine learning, the extracted features can be used to predict future sports dance movements accurately.

Sun et al. [15] has discussed the Deep learning-based approaches for emotional analysis of sports dance involve training a machine learning model or convolutional neural network (CNN) to detect and classify emotional cues in sports dance performance clips. The model is trained on a labeled clip dataset [26] and can detect and classify features such as facial expressions, body movements, and sound. After training, the model can accurately predict the emotional content of a given clip of a sports dance performance. This approach can be

used to gain insight into sports dance performances' emotional qualities and help improve their quality.

Lei et al. [16] has discussed the Reconstruction of physical dance teaching content and movement recognition based on a machine learning model uses the latest AI and machine learning technologies to automate the dance instruction process. Through powerful computer vision algorithms and specialized neural networks, the machines can be trained and utilized to detect, classify, interpret, and predict complex dance steps and choreographic sequences. This technology also enables the machine to recognize movements from both body and facial movements, making it possible to create innovative virtual dance instruction methods. The machine can accurately track a dancer's progress, providing feedback as he or she progresses through the program. The combination of computer vision and machine learning to create customized dance instruction methods allows for a more efficient and effective way of teaching.

Zhou et al. [17] has discussed the Research and Implementation of a Specific Action Generation Model in Dance Video Based on Deep Learning Technology (RSAGMDV-DLT) is an innovative solution to generate and track motion in video dance recordings using deep learning algorithms. This technology is capable of accurately predicting and determining the actions performed in a video dance recording without the need for manual annotation. The output is in the form of several poses and action files, which can be used for various applications, such as automated choreography and motion capture. The RAGMDV-DLT model can also enable real-time feedback and improvement of dance performances in various scenarios. This technology can potentially revolutionize how we capture, store, analyze, and reproduce dance videos.

Tang et al. [18] has discussed the Research on Dance Movement Evaluation methods based on Deep Learning Posture Estimation focuses on developing systems that use deep learning algorithms to detect and measure a dancer's movements. It can involve the analysis of photographs or videos of dance performances and real-time movement tracking. The goal is to accurately provide real-time insights into a dancer's performance, such as tempo, precision, and expression. In particular, deep learning posture estimation can be used to measure the quality of a dancer's movements, providing the ability to compare performances and offer feedback accordingly. It can be an efficient and accurate approach to analyzing a dancer's performance.

Zhang et al. [19] has discussed the Analyzing body changes of high-level dance movements through biological image visualization technology by a convolutional neural network (CNN) is a research methodology that uses computer vision and neural networks to analyze dancers' body movements. This type of research uses CNNs to detect changes in the body's position from a single image or sequence of images. The CNNs are used to identify patterns in the data and build models that can predict future body positions based on the patterns detected. This type of research could be used to improve techniques used in video analysis and provide insights into aspects of high-level dance movements, such as coordination and timing. Furthermore, this type of research could be used to create more expressive and accurate motion and motion capture animation.

Masurelle et al. [20] has discussed the Multimodal classification of dance movements using body joint trajectories and step sounds is an advanced form of machine learning, combining two separate forms of data—body joint trajectories (movement patterns captured by motion capture systems) and step sound recordings—to identify different dance movements. By combining the two different data sources, it is possible to create a model capable of recognizing a variety of complex movements. It provides the ability to detect and classify different types of dance, such as ballet and break-dance.

Shalini, A et al. [21] has discussed the Deep Learning of Dance Movements (DLPD) is a new approach to understanding and analyzing hip-hop dance movements from digital recordings. The goal of extracting key movement characteristics from such recordings is to understand better the nuances of hip-hop dance and how it is performed. Through DLPD, researchers can analyze the dynamics, body posture, and footwork of hip-hop dance to gain insights into the many variations of the dance form. Researchers hope to develop better tools for teaching and assessing hip-hop dance performance through this approach. DLPD can provide detailed insights into various parameters of the hip-hop dance movement. It can capture critical features such as dynamics (acceleration, speed, and deceleration), weight transitions, and body orientation. It can also identify body posture and style of movement, as well as movements' timing, complexity, and sharpness. Furthermore, DLPD can help detect external influences on a dancer, such as another dancer or the environment. It can help researchers gain insight into improvisations in each dancer's style and associated techniques.

Zhu et al. [22] has discussed the Dance Action Recognition and Pose Estimation based on Deep Convo-

lutional Neural Network (DCNN) uses a deep learning model to recognize dance actions and estimate pose from a sequence of frames. The deep learning model consists of a convolutional neural network (CNN) and a long short-term memory (LSTM) network. It can extract features from the input frames and use them to recognize the dance actions and estimate the dancers' poses. Furthermore, DCNN can recognize several levels of complexity within the dance, allowing for a more nuanced understanding of the dances.

Jin et al. [23] has discussed the One potential application of the fusion of deep learning biological image visualization technology and human-computer interaction intelligent robots in dance movements is in the creation of lifelike virtual dancers. By leveraging the power of deep learning technology, virtual dancers can be designed to imitate natural movements, learning from the movements of real people and interpreting them through various artificial intelligence models. Additionally, with the incorporation of robot-human interaction, the robots can be programmed to create smooth, lifelike motion, resulting in a level of realism that is impossible with traditional motion capture technology. This technology could have potential applications ranging from interactive dance games to AI-driven virtual choreography.

Lin et al. [24] has discussed the Dance movement recognition is an emerging technology that uses convolutional neural networks (CNNs) to recognize and interpret body movements from videos. The method allows for the automatic recognition of dance movements, including choreography, postures, and rhythms. It also enables the recognition of different styles and genres, including ballet, hip-hop, contemporary, and Latin. By combining the most advanced computer vision and machine learning techniques, the technology can detect and analyze posture, body shape, and rhythm to infer the most likely dance movement Dhiman G et al. [25]. These techniques enable detecting and classifying individual steps, creating automated choreographies, and enhancing the feedback system in dance classifications.

### 2.1. Research Gap.

1. Limited research on the robustness of the sensor-based Passive Dance Movement Recognition (PDMR) model across different forms of dance moves.
2. Incomplete understanding of how the IoT-based model can be used to differentiate between dance styles cost-effectively.
3. Very little research evaluates the success of IoT-based dance movement recognition models in large-scale datasets
4. Lack of reliable and effective machine learning models that accurately recognize dance motions from videos or images
5. There needs to be more research on the impact of wearable technology on the accuracy of the IoT-based dance movement recognition models
6. Minimal study on using low-cost and low-power sensors to enable efficient dance recognition

### 2.2. Research Objectives.

1. To develop a novel IoT-based dance movement recognition model based on a deep learning framework to identify the motion patterns and detect the motion features for dance movement recognition
2. To analyze the most accurate and efficient architecture for recognizing and classifying dance movements using deep learning and implement an IoT platform to integrate motion data collected from sensors with deep learning algorithms
3. To evaluate and compare the performance of the proposed system through an experiment involving dance movements of individuals and evaluate the accuracy of the proposed system in recognizing different types of dance movements
4. To analyze the efficiency and scalability of the proposed system and assess the usability and user experience of the proposed system

**3. Proposed Model.** The IoT-based dance movement recognition model based on a deep learning framework is a powerful system that leverages machine learning algorithms to recognize dance movements. It can recognize various dance movements, including African, Latin, and Contemporary styles. This model uses a camera and embedded sensors to recognize dance movements. It can be used to monitor the state of dancers, such as their speed, energy level, and range of motion. The model works by utilizing a deep learning framework that processes data captured from the sensors associated with the dancer's movements. This data then identifies

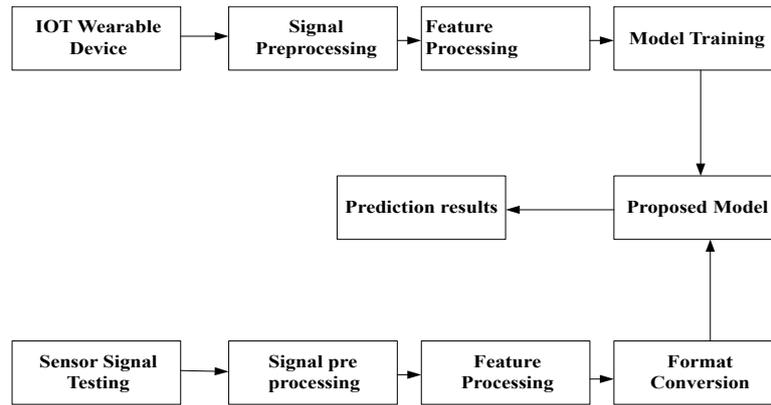


Fig. 3.1: Proposed block diagram

specific dance movements and generates statistical information. The block diagram of the proposed model has shown in the following Fig. 3.1.

This data is then used for identifying dance styles and providing feedback to the dancer. By recognizing individual dance movements, the model can provide feedback on improving their performance and allow the dancer to understand what works and what does not. This model can also be used for choreographing and creating dances, as it can recognize specific patterns that can be used to create unique choreographies

**3.1. Proposed Framework.** The IoT-based dance movement recognition framework is an intelligent system that tracks dancers' movements and generates motion-based visual feedback. It works by tracking a dancer's body movements and angles using a combination of IoT sensors, analyzing the real-time data to recognize specific dance moves, and using that information to create an interactive visual feedback loop. The framework also provides features such as music-timed animations, an easy way to create and share amongst friends, and the ability to store and share recorded dance performances. IoT technology in this framework provides an efficient and cost-effective way for dancers to receive objective feedback on dancing performances. In addition, motion-based visual feedback is an excellent tool for learning and developing new dance moves and techniques. The proposed framework has shown in Algorithm 1.

---

**Algorithm 1** IoT-based dance movement recognition framework

---

```

1: IP: A; // input IoT Signal;
2: Ma ( ); // dance motion capturing function;
3: Mm ( ); // dance appearance capturing function;
4: Tth ( ); // Tempering threshold function;
5: OP: KFL ( ); //key frame label;
6: Start
7: If  $|Mm ( ) > Ma ( )|$ 
8:  $SPLIT_{KFL} ( )$ ; // Split the key frame;
9: Else If  $|Mm ( ) < Ma ( )|$ 
10:  $JOIN_{KFL} ( )$ ; // Join the key frame;
11: Else If  $|Mm()^{Ma}()| = 0$ 
12:  $MOV_{KFL} ( )$ ; // Move the key frame;
13: Else If  $|Mm() = Ma()|$ 
14:  $STOP_{KFL} ( )$ ; // Stop the key frame;
15: COUNT =COUNT+1;
16: End
  
```

---

The IoT-based dance movement recognition framework is a system that utilizes Internet of Things (IoT) technology to recognize and analyze human motion during a dance performance. The system can track and identify the dancer's position and motion by sensing the dancer's movements through smart devices such as cameras, accelerometers, and gyroscopes. With this information, the system can provide quantitative analysis of the dancer's movements, generate automatic feedback for improvement or training purposes, and even recognize the style of the dancer's movements. Furthermore, this system can create applications such as score systems, data collection of dancers, and gesture recognition. Ultimately, the framework can improve dance training, performance, and evaluation accuracy and efficiency.

### 3.2. Dataset Description.

- The number of Dance moves are available in this dataset has 9 in dataset [25].
- The total number of available data in this dataset has 8467.
- The utilization of Training data (80%) is 6774 and testing data (20%) is 1693.

**3.3. Preprocessing.** Preprocessing is essential in developing a robust IoT-based dance movement recognition model based on a deep learning framework.

$$X(e|f) = \frac{X(e, f)}{X(f)} \quad (3.1)$$

Preprocessing is a critical stage in the development process as it prepares the data for training and later evaluation of the model.

$$X(e|f) = \frac{1}{X(f)} * \frac{1}{Y} \exp\{g^h f + g^d e + fRe\} \quad (3.2)$$

Preprocessing includes data cleaning, normalization, augmentation, feature extraction, dimensionality reduction, and feature selection.

$$X(e|f) = \frac{1}{Y'} \exp\{g^h e + fRe\} \quad (3.3)$$

These tasks are essential in preparing the data for the learning algorithms. Data cleaning helps improve the data quality by removing erroneous data points, outliers, and other meaningless values:

$$X(e|f) = \frac{1}{y'} \exp\left\{\sum_{p=1}^{q-e} g_p * e_p + \sum_{p=1}^{q_f} f_p R_p e_p\right\} \quad (3.4)$$

$$X(e|f) = \frac{1}{Y''} \pi_{p=1}^{q_f} \exp\{g_p * e_p + f_p R_p e_p\} \quad (3.5)$$

Normalization ensures that the data is within the same range and scales the data to improve learning accuracy. Data augmentation helps to improve the diversity of the data set by adding new samples with variations in existing data points.

**3.4. Feature Extraction.** Feature extraction is extracting meaning and relevant information from a raw data set. Dimensional reduction reduces the number of features while keeping most of the information intact.

$$X(e_q = 1/f) = \frac{X(e_q = 1/f)}{X(e_q = 0/f) + X(e_q = 1/f)} \quad (3.6)$$

Feature selection is choosing the most relevant features based on importance for the learning process.

$$X(g_p = 1/f) = \frac{\exp\{e_p + f^g H_{p,q}\}}{\exp\{0\} + \exp\{e_p + f^g H_{p,q}\}} \quad (3.7)$$

These steps help to reduce noise and computational costs and improve the learning accuracy of the model.

$$X(e_q = 1/f) = \beta(e_p + f^g H_{p,q}) \quad (3.8)$$

Its primary purpose is to reduce the dimensions of high-dimensional raw data to the feature dimensions. Feature extraction aims to represent important information from the raw data in a more straightforward and efficient form that deep learning algorithms can understand.

$$\frac{d}{de} \left( \frac{df}{de} \right) = \frac{d}{du} (E_g^{e^*} \text{Cos} E_g + g^e \text{sin} E_g) \quad (3.9)$$

$$\frac{d^2 f}{de^2} = \frac{d}{de} (E_g^a * \text{Cos} E_g) + \frac{d}{de} (g^e \text{sin} E_g) \quad (3.10)$$

$$\frac{d^2 f}{de^2} = g^e \frac{d}{de} (\text{Sin} E_g) + \text{Sin} E - g \frac{d}{de} g^e + E_g^e \frac{d}{de} (\text{Cos} E_g) + \text{Cos} E_g \frac{d(E_g^e)}{de} \quad (3.11)$$

In the context of a dance movement recognition model, feature extraction can be used to identify the movements and patterns of a dancer by extracting distinct features related to body orientation, relationship data between body joints, and more.

$$\frac{d^2 g}{de^2} = E_g^a \text{Cos} E_g - E^2 g^e \text{Sin} E_g + g^e \text{Sin} E_g + E g^e \text{Cos} E_g \quad (3.12)$$

$$\frac{d^2 f}{de^2} = 2 E g^e \text{Cos} E_g - E^2 g^e \text{Sin} E_g \quad (3.13)$$

By extracting these features, the model can analyze the data to recognize different types of dance movements. In addition, feature extraction also helps reduce computational workload, allowing the algorithm to process large amounts of data quickly and accurately.

**3.5. Dance Movement Detection.** The functions of Dance Movement Detection in an IoT-Based Dance Movement Recognition Model based on Deep Learning Framework can be classified into two main categories—detection, recognition, and classification. The dance moment detection has shown in the following Fig. 3.2.

$$\frac{df}{de} = \lim_{g \rightarrow 0} \frac{g(e+f) - g(e)}{h} \quad (3.14)$$

The first category is the detection of dance movements. It involves detecting the features of a dancer's body posture and movements and transforming them into high-level digital representations:

$$\frac{df}{de} = \lim_{g \rightarrow 0} \frac{1/(e+f) - 1/e}{h} \quad (3.15)$$

This kind of feature detection is necessary to identify different types of dance movements. It uses basic image processing or more advanced systems such as convolutional neural networks. The second category is the recognition and classification of dance movements.

$$\frac{df}{de} = \lim_{g \rightarrow 0} \frac{\frac{1}{e+f} * \frac{e}{e} - \left( \frac{1}{e} * \frac{(e+f)}{(e+f)} \right)}{h} \quad (3.16)$$

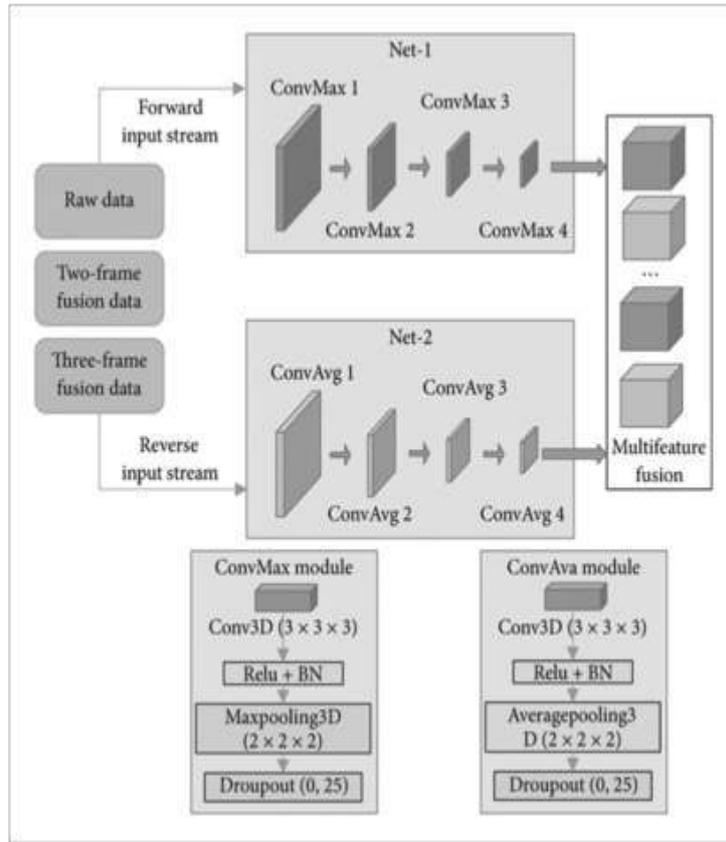


Fig. 3.2: Detection of dance movements

$$\frac{df}{de} = \lim_{g \rightarrow 0} \frac{\frac{e-e-g}{(e+g)*e}}{h} \tag{3.17}$$

It involves identifying the type and form of the dance movement from the digital representations. It is usually done using machine learning techniques such as decision trees, support vector machines, neural networks, and deep learning models.

$$\frac{df}{de} = \lim_{g \rightarrow 0} \frac{\frac{-g}{(e+g)*e}}{h} \tag{3.18}$$

$$\frac{df}{de} = \lim_{g \rightarrow 0} \frac{\frac{-1}{(e+f)*e}}{h} \tag{3.19}$$

$$f = \frac{-1}{e^2} \tag{3.20}$$

These models can be trained to classify the dance movements into different forms, such as salsa dance, freestyle, etc. Overall, Dance Movement Detection in an IoT-based Dance Movement Recognition Model based on Deep Learning can detect various dance movements and accurately classify them into different types. It also enables real-time performance enhancement given good-quality input data.

To improve the generalization of the deep learning framework used for the given task, one can employ a few techniques such as:

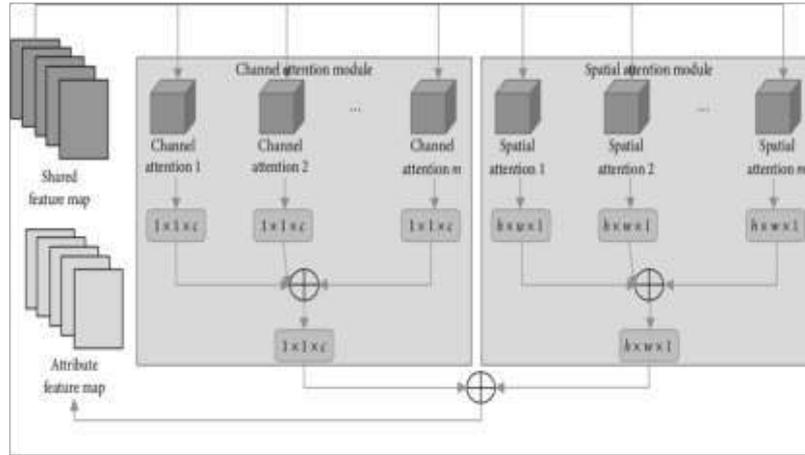


Fig. 3.3: Detection of dance movements

1. Regularization: Regularization is a very common technique used in machine learning to reduce generalization error and prevent model overfitting by introducing a penalty term which encourages the model to generate less complex decision boundaries. It helps regularize the model to ensure that it is able to learn from fewer data points and generalize the patterns learned. Popular regularization techniques for deep learning models include weight regularization (L1 & L2), Dropout, Max-norm constraints, Data Augmentation, Early Stopping, etc.
2. Using a smaller neural network: Using a smaller network with fewer layers, nodes, and parameters helps avoid over-fitting to the training data and promotes generalization capabilities of the model.
3. Transfer learning: In transfer learning, one can take a pre-trained deep learning model from a believed source or use a known library of models and fine-tune on the data for the given task. This technique also helps in better generalization of the deep learning framework.

**3.6. Dance movement classification.** The function of Dance Movement Classification in an IoT-based dance movement recognition model based on a deep learning framework is to accurately recognize and categorize specific movement patterns captured by connected devices.

$$G = \lim_{e \rightarrow 0} \frac{g(f+e) - g(f)}{e} \quad (3.21)$$

$$G = \lim_{e \rightarrow 0} \frac{\frac{1}{f+e-1} - \frac{1}{f-1}}{e} \quad (3.22)$$

The deep learning model can be used to classify different movements in a dance form, such as hip-hop, break-dance, ballet, rhythm tap, and so on. The classification of dance movements has shown in the following Fig. 3.3.

$$G = \lim_{e \rightarrow 0} \frac{\frac{1}{f+e-1} \frac{f-1}{f-1} - \frac{1}{f-1} \frac{f+e-1}{f+e-1}}{e} \quad (3.23)$$

$$G = \lim_{e \rightarrow 0} \frac{(f-1) - (f+e-1)}{e(f-1)(f+e-1)} \quad (3.24)$$

$$G = \lim_{e \rightarrow 0} \frac{-e}{e(f-1) * (f+e-1)} \quad (3.25)$$

Additionally, the model can learn to differentiate between different techniques professionals and amateur dancers use. It can then be used to group the users according to their dancing abilities and create a database of their performances.

$$G = \lim_{e \rightarrow 0} \frac{(-1)}{(f-1) * (f+e-1)} \quad (3.26)$$

$$G = \frac{(-1)}{(f-e)^2} \quad (3.27)$$

Furthermore, the deep learning model can also be used to identify and record changes in movement patterns over time, thus allowing for high accuracy when analyzing the user's dancing abilities.

1. *Efficiency.* Optimizing the model for efficiency can be done through techniques such as model pruning, data compression, and alternative network architectures. In addition, using a distributed computing architecture can also help speed up the computation of the model.

2. *Precision.* Measuring the performance of the model is crucial to ensure it is working with the desired level of accuracy. It can be done by performing comparisons between the model output and ground truth labels. Additionally, augmenting the training data with additional samples of hard cases can help to increase the model's robustness and accuracy.

3. *Scalability.* Making the model scalable for use in real-world applications requires implementing the model in an efficient and robust software environment. It includes technologies such as containerization and cloud computing, which provide low-cost computing access and low-latency performance. Additionally, deploying the model using different software frameworks can help manage its scalability.

4. *Interoperability.* Enabling the model to understand contextual cues and analyze complex motion patterns requires interoperable software architecture. It might include connecting to external services such as speech recognition APIs or setting up interfaces to integrate with other software systems easily. In addition, this can also include setting up an API to provide access to the model's data and predictions.

Deep learning frameworks enable developers to quickly design and develop deep learning models with simple and intuitive interfaces. Depending on the type of model and the data requirements, one framework may offer better results than the other. Other frameworks may also be used depending on the specific requirements of the model. In order to account for data latency in the IoT network, the deep learning framework used should support an ongoing learning process that can account for changes in data over time. It will keep the model up to date without having to manually re-train it. The framework should provide built-in security measures such as encryption, authentication, and access control. The framework should include comprehensive security features that can protect sensitive data from being accessed or tampered with maliciously.

*Utilize Confusion Matrix and Classification Report:* A confusion matrix is a table showing different predictions made versus the true values and can provide insight into where the model is going wrong. The classification report is a summary of the results, including precision, recall, and F1-score.

*ROC-AUC Curve Analysis:* A Receiver Operating Characteristic (ROC) curve is a graphical representation of a model's true positive rate (TPR) against its false positive rate (FPR) at various threshold values. AUC stands for Area Under the Curve and is a metric that quantifies model performance.

*Stratified K-fold Cross Validation:* Stratified K-fold cross-validation can be used to address class imbalance issues by ensuring that each fold is balanced. This method allows the model to be tested on a range of different training/test splits to ensure that the model's performance is consistent over the whole dataset.

*Evaluate the Model on Unseen Data:* To make sure that the results of the model can be used with confidence by stakeholders, evaluate it on unseen data using the same techniques used to develop the model, such as evaluating the AUC, accuracy, and F1-score. It will ensure that the model is generalizing well to unseen data.

*Estimate the Model's Performance Over Time:* Estimate the model's performance over time by testing it on data from different periods and evaluating changing performance metrics such as accuracy and F1-score. It will help identify any long-term issues with the model's performance.

Table 4.1: Computation of Accuracy (in %)

No.of Inputs	2DMCM	DAGM	DMDL	EDRM	IDLF
100	47.41	55.06	60.74	74.21	85.82
200	48.90	57.03	63.16	76.41	87.81
300	49.70	58.16	63.57	77.21	89.01
400	50.96	59.85	65.32	78.94	90.74
500	52.10	61.40	66.73	80.44	92.33
600	53.25	62.95	68.15	81.94	93.93
700	54.39	64.50	69.56	83.44	95.53

*Feature Importance Analysis:* Feature importance analysis will help identify which features are having the most impact on the model's performance. It can be used to refine the model and to make sure that the most important features are being used.

**4. Comparative Analysis.** The proposed IoT deep learning framework (IDLF) has compared with the existing two-dimensional matrix computation model (2DMCM), Dance Action Generation Model (DAGM), Dance Movement Based on Deep Learning (DMDL) and Edge Distance Random Matrix (EDRM). Here, python is a simulation tool used to execute the results.

Accuracy and reliability can be assessed using a number of metrics such as precision, recall, and F1-score. The model must be trained on large datasets. The datasets need to contain different dancers with different poses. The model should be tested for performance on different poses. It is also important to ensure that the data used for training and testing is diverse with respect to the poses, angles of each pose, different lighting, and different locations. It is important to ensure the model is reliable and accurate when deployed on large datasets. The model should be trained on datasets that include poses that are similar to those the model will be used for. It will help the model to generalize better and ensure it can recognize different poses of dancers.

Along with assessing the accuracy and reliability of the model, it is also important to assess the usability and interpretability of the proposed model. It can be done with an evaluation process that tests the application's user interface, the accuracy of the predictions, and the quality of the accuracy measures. Moreover, an evaluation system should be established to test the model's performance on different tasks, such as identifying different types of dancers, different poses, different features in the poses, and different transitions between the poses. It will help to make sure the model is capable of recognizing and mapping different poses of dancers.

*Update dataset:* It is important to update the dataset with the latest available data regularly to make sure the model uses the most up-to-date data. It should be done as often as the data changes and new data is available.

*Evaluate hardware requirements:* The model should be tested with multiple hardware configurations to ensure that it can accommodate the hardware requirements of the deployed system without adversely affecting accuracy or performance.

*Constraints:* All relevant constraints should be considered when designing the model. It includes limitations on memory, time, compute power, energy, etc. Any limitations should be addressed in the development and deployment of the model, and solutions should be found to accommodate the constraints.

*Deployment:* Make sure the model can be deployed with minimal changes to existing infrastructure. It requires planning upfront to ensure that the model can be deployed without disrupting existing operations and that all necessary dependencies are accounted for.

**4.1. Computation of Accuracy.** The accuracy of an IoT-based dance movement recognition model based on a deep learning framework can be computed by first training the model with a dataset of labeled dance movements. The accuracy is then computed by running the trained model on a test set of unlabeled movement data and comparing the predicted labeling results with the actual labels for each data point. It measures how accurately the model can recognize and differentiate between different types of dance movements. Table 4.1 shows the computation of accuracy.

Fig. 4.1 shows the computation of accuracy. A high accuracy indicates that the model can correctly classify

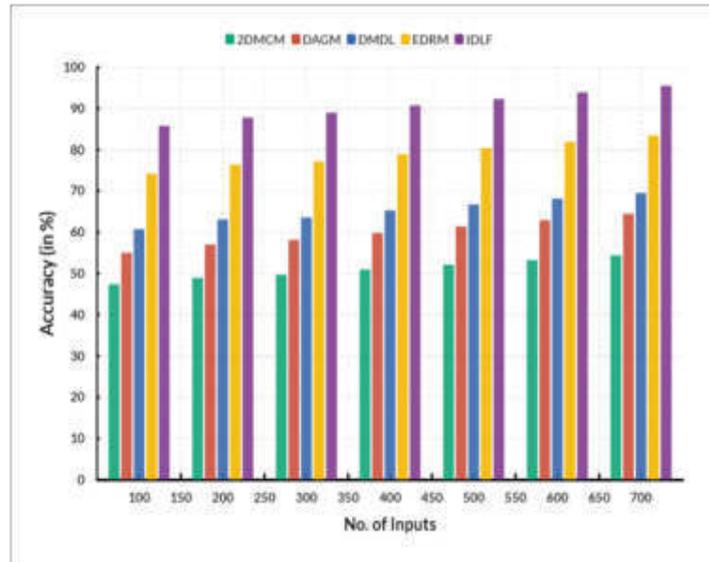


Fig. 4.1: Detection of dance movements

Table 4.2: Computation of Precision (in %)

No. of Inputs	2DMCM	DAGM	DMDL	EDRM	IDLF
100	52.11	61.40	66.74	80.44	82.33
200	53.25	62.95	68.15	81.94	83.93
300	54.40	64.50	69.57	83.44	85.52
400	55.54	66.05	70.98	84.94	87.12
500	56.69	67.60	72.40	86.44	88.71
600	57.83	69.15	73.81	87.94	90.31
700	58.98	70.70	75.23	89.44	91.90

the movements and identify their distinguishing characteristics. A model with a lower accuracy would likely need to accurately differentiate between two similar movements or label a given movement.

The accuracy of feature extraction done on the input dance movement can be checked by running a series of tests on the model, including comparing the results to real-world datasets. It can be done by comparing the extracted features of the input dance movement to the features of datasets of real-world dance styles. It will help to identify any potential errors or inaccuracies in the model's feature extraction. Once the accuracy of the model has been established, it can then be tested on different real-world datasets to assess how accurate the model is at detecting different dance styles. This type of testing can also be used to refine the model by adjusting its parameters or feature extraction algorithm. Ultimately, testing the model on different real-world datasets is the best way to determine its accuracy in detecting different dance styles.

**4.2. Computation of Precision.** Precision in an IoT-based dance movement recognition model based on a deep learning framework is the measure of correctness for the model's predictions. It is computed as the ratio of true positives (correctly predicted dance movements) and the total number of predicted dances. It measures the model's accuracy in correctly identifying dance movements from a particular dataset. Table 4.2 shows the computation of Precision.

Fig. 4.2 shows the computation of Precision. The precision can be improved by increasing the dataset size and making the model's layers deeper, increasing the number of predicted accurate positive results. Additionally, regularization parameters such as dropout, batch normalization, and weight initialization may be used to reduce

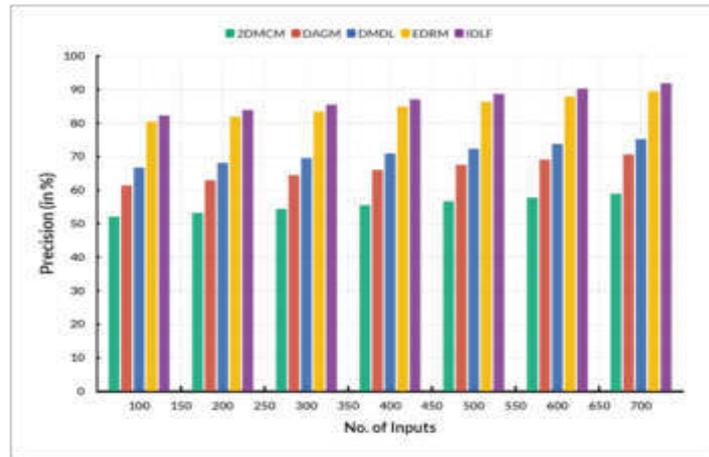


Fig. 4.2: Precision

Table 4.3: Computation of Recall (in %)

No. of Inputs	2DMCM	DAGM	DMDL	EDRM	IDLF
100	52.00	59.29	65.17	77.88	79.49
200	53.04	59.74	67.49	79.31	80.92
300	54.08	60.19	69.81	80.74	82.35
400	55.12	60.64	72.13	82.17	83.78
500	56.16	61.09	74.45	83.60	85.21
600	57.20	61.54	76.77	85.03	86.64
700	58.24	61.99	79.09	86.46	88.07

the cost of predicting false positives and increase model efficiency.

**4.3. Computation of Recall.** Recall is a measure of model performance used to evaluate how many relevant items a model successfully identified. In the context of a dance movement recognition model based on a deep learning framework, recall is computed by comparing the model's predicted output against the actual labels. Table 4.3 shows the computation of Recall.

Fig. 4.3 shows the computation of Recall. The recall is then calculated as the ratio between the number of correctly identified items and the total number of items in the data set. For example, if the model predicts movements for ten dances and seven are correctly identified, while three are not, the recall is 70%. It indicates that the model successfully recalled 70% of the accurate labels present in the dataset.

**4.4. Computation of F1-Score.** The F1-Score is a metric to evaluate a model's performance in classifying data into different classes. This measure is used in Machine Learning when dealing with classification problems, such as in the case of a Dance Movement Recognition (DMR) model based on a Deep Learning framework. Table 4.4 shows the computation of F1-Score.

Fig. 4.4 shows the computation of F1-Score. The F1-Score is calculated by considering the Precision and Recall of a model's performance. Precision refers to the model's ability to identify all of a class's instances accurately. At the same time, recall is the model's ability to detect each class instance that it is supposed to detect. The F1-Score is then calculated by taking the precision and recall values' harmonic mean (also known as the F1-Score). It is done by adding the precision and recall values and dividing them by two.

**5. Conclusion.** The IoT-based dance movement recognition model based on a deep learning framework effectively recognizes dance movements from sensors placed in a dancer's clothing or on the floor. By combining

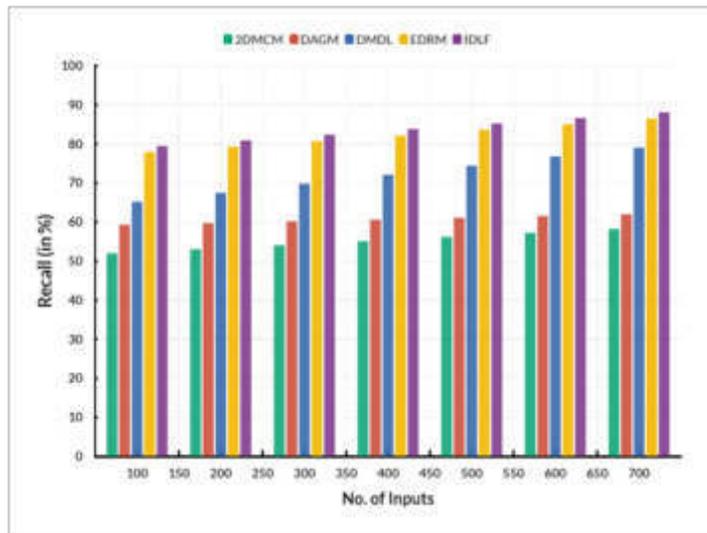


Fig. 4.3: Recall

Table 4.4: Computation of F1-Score (in %)

No. of Inputs	2DMCM	DAGM	DMDL	EDRM	IDLF
100	53.68	61.19	61.74	80.40	79.08
200	54.34	61.67	64.47	80.88	80.85
300	55.00	62.15	67.20	81.36	82.62
400	55.66	62.63	69.93	81.84	84.39
500	56.32	63.11	72.66	82.32	86.16
600	56.98	63.59	75.39	82.80	87.93
700	57.64	64.07	78.12	83.28	89.70

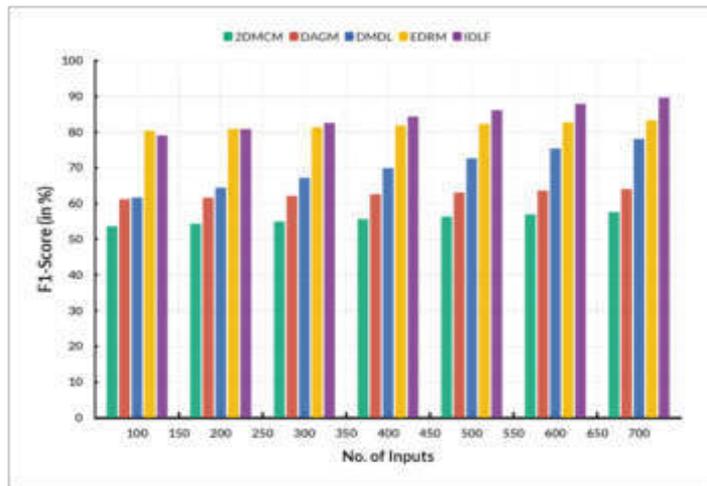


Fig. 4.4: F1score

the data from the sensors with deep learning algorithms, the model can be trained to recognize specific dances and movements with very high accuracy. Additionally, deep learning algorithms provide the flexibility to continuously adjust the model according to new dance moves and patterns. With the help of edge computing, the model can be deployed quickly and with limited resources. It is an effective tool for recognizing dance movements and can be used in various applications. The proposed model reached 90.74% accuracy, 87.12% precision, 83.78% recall and 84.39% F1-Score. The future scope of an IoT-based dance movement recognition model based on a deep learning framework is extensive. As the scope of AI and deep learning expands, IoT applications can be developed for many applications, including dance movement recognition. These applications can be designed to detect and recognize a range of dances and analyze how the body moves through different dance moves. Furthermore, IoT applications can collect and analyze data from dancers and incorporate them into a virtual environment for training and evaluation. It will allow dancers to observe and analyze how their body performs during movements, giving them valuable feedback on improving their dance style. Additionally, AI and deep learning can help dancers better understand their dancing habits and improve posture and style for better and more comfortable dance performances. Through this application, dancers and choreographers can create better performances and have a better understanding and appreciation for the art of dance.

## REFERENCES

- [1] Zhang, Y., & Zhang, M. (2021). Machine learning model-based two-dimensional matrix computation model for human motion and dance recovery. *Complex & Intelligent Systems*, 7, 1805-1815.
- [2] Ma, X., & Wang, K. (2022). Dance Action Generation Model Based on Recurrent Neural Network. *Mathematical Problems in Engineering*, 2022.
- [3] Feng, H., Zhao, X., & Zhang, X. (2022). Automatic Arrangement of Sports Dance Movement Based on Deep Learning. *Computational Intelligence and Neuroscience*, 2022.
- [4] Liu, X., & Ko, Y. C. (2022). The use of deep learning technology in dance movement generation. *Frontiers in Neurorobotics*, 16, 911469.
- [5] Jin, Y. (2022). A Three-Dimensional Animation Character Dance Movement Model Based on the Edge Distance Random Matrix. *Mathematical Problems in Engineering*, 2022.
- [6] Huang, Y. (2022). Comparative Analysis of Aesthetic Emotion of Dance Movement: A Deep Learning Based Approach. *Computational Intelligence and Neuroscience*, 2022.
- [7] Yuan, X., & Pan, P. (2022). Research on the Evaluation Model of Dance Movement Recognition and Automatic Generation Based on Long Short-Term Memory. *Mathematical Problems in Engineering*, 2022.
- [8] Ling, X. (2022, December). Movement Tracking Detection of Break Dance Based on Deep Learning. In *2022 6th Asian Conference on Artificial Intelligence Technology (ACAIT)* (pp. 1-5). IEEE.
- [9] Wang, S., & Tong, S. (2022). Analysis of high-level dance movements under deep learning and internet of things. *The Journal of Supercomputing*, 78(12), 14294-14316.
- [10] Yan, M., & He, Z. (2022). Dance Action Recognition Model Using Deep Learning Network in Streaming Media Environment. *Journal of Environmental and Public Health*, 2022.
- [11] Ahmed, S. T., Kumar, V. V., & Kim, J. (2023). AITel: eHealth Augmented-Intelligence-Based Telemedicine Resource Recommendation Framework for IoT Devices in Smart Cities. *IEEE Internet of Things Journal*, 10(21), 18461-18468. <https://doi.org/10.1109/jiot.2023.3243784>.
- [12] Zhang, Y. (2022). Application of Knowledge Model in Dance Teaching Based on Wearable Device Based on Deep Learning. *Mobile Information Systems*, 2022.
- [13] Zhang, A. (2023). Optimization Simulation of Match between Technical Actions and Music of National Dance Based on Deep Learning. *Mobile Information Systems*, 2023.
- [14] Karthick Raghunath, K. M., Koti, M. S., Sivakami, R., Vinoth Kumar, V., NagaJyothi, G., & Muthukumaran, V. (2022). Utilization of IoT-assisted computational strategies in wireless sensor networks for smart infrastructure management. *International Journal of System Assurance Engineering and Management*. <https://doi.org/10.1007/s13198-021-01585-y>.
- [15] Lei, L., & Tingting, Y. (2023). Reconstruction of physical dance teaching content and movement recognition based on a machine learning model. *3 c TIC: cuadernos de desarrollo aplicados a las TIC*, 12(1), 267-285.
- [16] Zhou, L. (2021, December). Research and Implementation of Specific Action Generation Model in Dance Video Based on Deep Learning Technology. In *2021 International Symposium on Advances in Informatics, Electronics and Education (ISAIEE)* (pp. 243-247). IEEE.
- [17] Tang, H., Luo, Y., & Yang, J. (2022, August). Research on Dance Movement Evaluation Method Based on Deep Learning Posture Estimation. In *2022 2nd International Conference on Big Data Engineering and Education (BDEE)* (pp. 173-177). IEEE.
- [18] Zhang, R. (2022). Analyzing body changes of high-level dance movements through biological image visualization technology by convolutional neural network. *The Journal of Supercomputing*, 78(8), 10521-10541.
- [19] Masurelle, A., Essid, S., & Richard, G. (2013, July). Multimodal classification of dance movements using body joint trajectories and step sounds. In *2013 14th international workshop on image analysis for multimedia interactive services (WIAMIS)*

- (pp. 1-4). IEEE.
- [20] Shalini, A., Jayasuruthi, L., & VinothKumar, V. (2018). Voice recognition robot control using android device. *Journal of Computational and Theoretical Nanoscience*, 15(6-7), 2197-2201.
  - [21] Zhu, F., & Zhu, R. (2021). Dance Action Recognition and Pose Estimation Based on Deep Convolutional Neural Network. *Traitement du Signal*, 38(2).
  - [22] Jin, N., Wen, L., & Xie, K. (2022). The fusion application of deep learning biological image visualization technology and human-computer interaction intelligent robot in dance movements. *Computational Intelligence and Neuroscience*, 2022.
  - [23] Lin, Z. (2023, May). Dance movement recognition method based on convolutional neural network. In *2023 4th International Conference on Computer Vision, Image and Deep Learning (CVIDL)* (pp. 255-258). IEEE.
  - [24] Dhiman, G., Vinoth Kumar, V., Kaur, A., & Sharma, A. (2021). Don: deep learning and optimization-based framework for detection of novel coronavirus disease using x-ray images. *Interdisciplinary Sciences: Computational Life Sciences*, 13, 260-272.
  - [25] Dataset:<https://www.kaggle.com/datasets/nwjbrandon/human-activity-recognition-dancing>

*Edited by:* Polinpapilinho Katina

*Special issue on:* Scalable Dew Computing for Future Generation IoT systems

*Received:* Sep 30, 2023

*Accepted:* Dec 13, 2023



## RESEARCH ON NETWORK SECURITY SITUATION AWARENESS TECHNOLOGY BASED ON SECURITY INTELLIGENT MONITORING TECHNOLOGY

BINGYU YANG\*

**Abstract.** This paper uses data mining technology to dynamically monitor tobacco Industrial Enterprise' information systems. This paper builds an Internet security situation awareness system under a big data environment. The weight clustering method is used to classify users' network behavior. The spacing of weights is optimized to ensure the maximum difference in classification. Then, NAWL-ILSTM technology establishes a security situational awareness model for the Internet environment. In this project, the extended and short-memory Nadam optimal algorithm (NAWL) is used to realize data deep learning. Finally, the tobacco industry network security situation assessment method is designed to complete the dynamic monitoring of tobacco industry network security based on data mining. Simulation results show that the proposed method can effectively improve the safety evaluation performance of the system and reduce evaluation errors.

**Key words:** Association analysis; Intelligent monitoring; Algorithm detection; Situational awareness; Network security

**1. Introduction.** The US Air Force developed situational awareness technology in the 1980s. Its role is to analyze and judge the situation of the war and provide relevant information for the situation of the war. It is used in high-level decisions to win wars. At the network security level, situation perception research focuses on the analysis of information about the forms and development trends of network attacks. The research of situation awareness has the characteristics of global, dynamic, complex, effective and accurate. The research on situation awareness can be divided into three levels [1]. The first level of situational awareness refers to collecting large amounts of information or data. The main content includes host, network, security, application, physical, intelligence, threat, etc. The second level is to discuss the association and data integration. Data fusion is a method to process multiple observations from a timing sequence on a computer. It includes data association, merging, and extraction. It displays assets, threats, risks, weaknesses and trends at all levels [2]. The visualization, isomerization, automation and real-time processing of multi-source information are realized through the in-depth study of situation awareness. It provides a scientific basis for risk assessment, decision making and prediction. In recent years, data mining technology has been applied more and more. Some scholars have improved the MD5 algorithm and carried on the hardware design. Although this method can solve hardware-related problems such as encryption and decryption of situational awareness, there is still a lack of targeted methods for accurate authentication in user clustering. Currently, the existing hierarchical and clustered WSN protocol only carries out the optimal verification for different types of user groups and lacks anything related to hardware. Applying the existing theory and technology to practice is a complex problem. Therefore, this paper uses data mining technology to realize the situation awareness and security monitoring of tobacco enterprise information systems.

### 2. Network security situation awareness system structure.

**2.1. Overall system framework structure.** Figure 2.1 shows the architecture of the network security situational awareness system. The architecture is distributed and open architecture. Its core concepts are "decentralized access" and "partition management". It can be divided into three levels: information acquisition, factor extraction, and scenario decision-making. This system uses a log-based sensor, SNMP sensor and Net-Flow sensor to collect data [3]. The network environment's host, switching equipment, and security equipment are analyzed and studied. The element extraction layer is designed for obtaining many different types of information. The compression and extraction of information are realized by employing aggregation and fusion. The

---

\*China Tobacco Anhui Industrial Co., LTD., Hefei, 230088, China ([magicby123@163.com](mailto:magicby123@163.com))

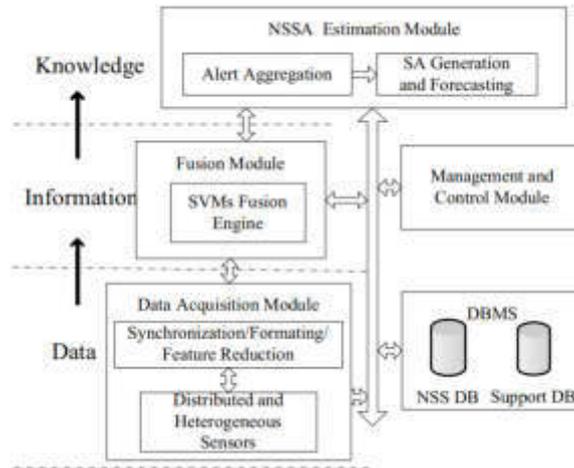


Fig. 2.1: Network security situational awareness system architecture.

situation decision-making level realizes the comprehensive cognition of multi-source information using hierarchical evaluation and the dynamic cognition of situation awareness using nonlinear time series prediction. It gives high-level users an intuitive understanding of the system's security status. In implementing the information acquisition layer, factor extraction layer and scene decision layer, it is necessary to cooperate with the relevant database. It includes an event database, resource database, network information database, knowledge base, etc. Building such a database requires the assistance of experts, security personnel and network scans.

An information perception method based on "monitor-analysis-decision" is proposed. A variety of sensors are used for safety monitoring. Each sensor node can collect corresponding information to achieve the monitoring of the overall operation of the system. Then this paper adopts the information filtering method, verification and fusion to realize the security situation awareness of the global state.

**2.2. System Physical Architecture.** FIG. 2.2 is a schematic diagram of the system's physical architecture described in this paper. The system includes a sensor, analyzer, decision maker and corresponding database. In each safe zone, there are multiple sensors and multiple analysis programs [4]. In addition to the detection and analysis units, the safety area also contains a discrimination unit. The sensor monitors the central system and local area network and transmits abnormal and suspicious activity reports to the analyzer. The analysis program mainly obtains information from local sensors. And transfer that data to a global database. Decision makers obtain the overall network security degree by analyzing each safety zone and storing it in the database. The two sensors are connected circularly. The goal is to complement information across multiple sensors so that each sensor perceives the raw data. This simplifies the raw data. Similar links are also used in the various analysis programs to achieve better precision in element extraction and consistent publication of the overall strategy. In the process of data acquisition, the analysis instrument and the detection instrument adopt a two-way interactive way. On the one hand, the sensor will actively feed the collected data to the analytical machine. On the other hand, the analyzer can also issue a command to reconstruct the sensor. The analysis results of the local analyzers must not only be stored locally but also transmitted via encrypted channels to the database of the remote central monitoring system. It is submitted to the decision machine along with the results of the resulting global analysis.

Given the wide distribution of large-scale internet, many nodes, operating system and network device differences, multi-level information fusion and decision-making technology are needed. The global, real-time and dynamic security state perception of the network uses multi-source and heterogeneous information. It has the characteristics of self-adaptation and self-learning [5]. It reflects the characteristics of distribution and dynamics. It has both dynamic and static network attack detection functions. This architecture has nothing to do with network topology. This architecture can dynamically adjust the system according to different distributed

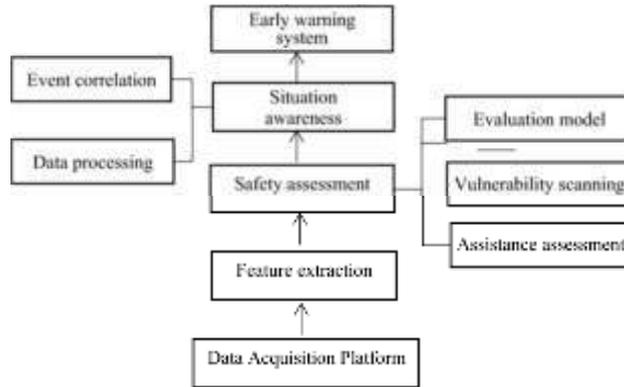


Fig. 2.2: Network security situation awareness model in multi-sensor network environment.

applications and security requirements. In addition, the architecture can effectively solve the problem of a single point of failure and improve the flexibility of the whole system [6]. At the same time, organizations within each security domain still adopt a hierarchical structure, significantly reducing the implementation complexity. This method provides a practical information acquisition method for security situation awareness in distributed large-scale networks.

**2.3. System Conceptual Architecture.** This paper presents the basic architecture of network security situation awareness. It is mainly used to reflect the flow of information at various levels of perception. From the lowest point of view to the highest point of view are the network security situation element extraction, situation assessment, and situation prediction. It is matched by data information, feature information, and situation information. The first layer is "network security scenario element extraction," which is the basis for the cognition of network security scenarios. At this level, we will focus on extracting the information that can play an essential role in future network security from a large number of multi-source security data. And convert it into a standard XML file. The research results of this project will provide necessary technical support for information security evaluation and early warning in the future network environment of our country [7]. The second layer is network security situation assessment, which is the core of network security situation awareness. It analyzes data from various scenarios to calculate possible hazards to services, hosts, and networks. Level 3 is the prediction of the network security situation. The system predicts the future network security situation according to the past and present network security situation. It allows policymakers to control potential dangers in advance to make the right decisions (Fig.2.3).

**3. Design of network security situation awareness algorithm.**

**3.1. Adaptive weighted clustering method.** A dynamic clustering model based on weights is proposed. It eliminates the distinction between data characteristics. Then, a method based on adaptive weighted clustering is proposed to fuse the features with the same network behavior [8]. The method aims to minimize the sum of squares of inter-class deviations. The guiding principle based on this is to maximize the differences between groups. Select a small number of feature sets according to the principle of random non-repetition. Finally, the solution of the increased in-class variance is obtained. The sum of squares of the errors of each cluster determines the weighting of each cluster. When a physical object is reassigned, it must be set according to the distance of the weight. The minimization algorithm with weights is adopted to ensure the maximum difference in classification [9]. The proposed method can effectively improve the system's running speed and reduce the system's calculation amount to achieve the optimal purpose. This method can not only overcome the dependence of traditional clustering methods on cluster center selection but also effectively solve the problem of large-scale feature sets. The flow of this algorithm is shown in Figure 3.1.

(1) Characteristics of standardization. Use  $U = \{u_1, u_2, \dots, u_n\}$  to represent the set of all IP pairs of communication example features.  $n$  represents the number of communication instances of IP.  $S = \{s_1, s_2, \dots, s_m\}$

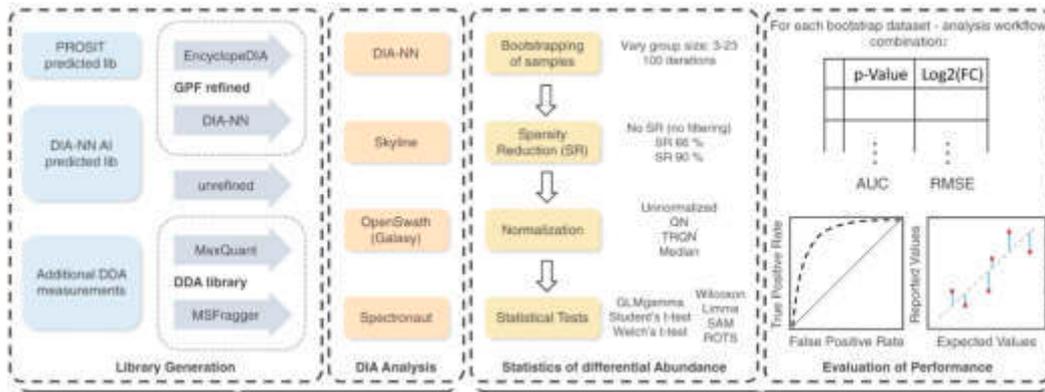


Fig. 2.3: Conceptual structure of network security situation awareness.

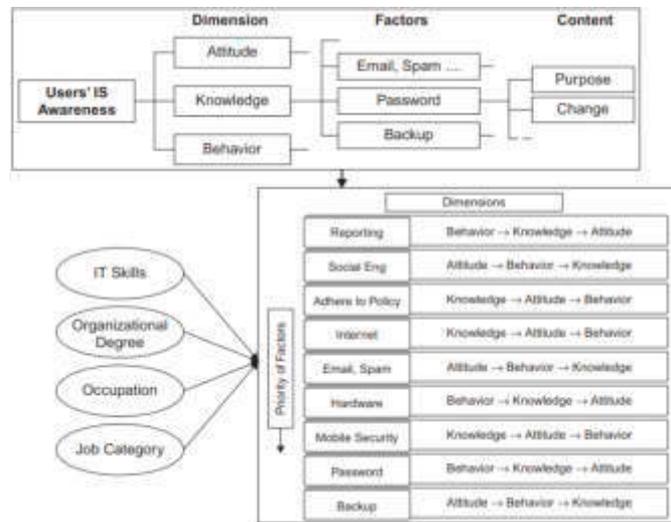


Fig. 3.1: Network in-depth behavior analysis process.

stands for a set of cluster centers.  $B_i = \{u_{i,j} | 1 \leq j \leq m\}$ .  $B_i$  is represented by a set of eigenvectors. The  $i$  example uses the representation of  $B_i$ .  $u_{i,j}$  represents the  $j$  function under  $i$  of the communication. It normalizes the eigenvalues in the range 0 to 1.

(2) The weighted clustering method is adopted when the center of cluster  $m$  is any initial value.

(3)  $Z$  similar method is used to perform weighted clustering when determining the dimensions in the feature subset for which  $A$  is initialized.

(4) Select any one of  $B^l = \{u_i\}_{i=1}^Z$  number of attribute examples in the group.

(5) Calculate the distance between the cluster center and each feature example of the feature subset in each feature example. The samples are classified according to the shortest weight [10]. In this paper, an adaptive clustering method based on weighting is proposed to optimize the weighting of each index. By adaptive adjustment to a cluster, the weight of each cluster is adjusted, and the alternation of the largest and smallest

steps is carried out. The calculation formula to obtain weighting  $\sigma_j$  is (3.1).

$$\sigma_j = \eta \sigma_j^{t-1} + (1 - \eta) \left( \frac{R_j^{\frac{1}{1-\alpha}}}{\sum_{j'=1}^M R_{j'}^{\frac{1}{1-\alpha}}} \right), 0 \leq \eta \leq 1 \quad (3.1)$$

$\sigma_j^{t-1}$  stands for weighting coefficient.  $\eta$  represents the control factor that is weighted repeatedly.  $R_j^{\frac{1}{1-\alpha}}$  stands for different weighting characteristics. The current update is determined by the previous iteration weight  $\eta$ . The algorithm can ensure smooth transformation of weighted values during iteration and maintain them within the maximum weighted range during iteration [11]. The premise of minimizing the sum of squared errors is to achieve higher clustering weights. The initial value  $\alpha_{init}$  is equal to 0. In each iteration step, the  $\alpha_{step}$  value gradually increases to  $\alpha_{init}$  until it reaches the maximum value  $\alpha_{max}$ . After the maximum value  $\alpha_{max}$ , the morphology of the cluster remained at a stable level.

(6) For the cluster center, it needs to be re-calculated.

(7) Return to step (4) to perform the calculation repeatedly. Until the cluster center does not change or the maximum number of iterations is reached.

(8) Calculate the sum of squares of error of the new class  $M$  to get the difference value  $R_\alpha$ .

(9) If  $R_\alpha \geq R_{max}$  exists,  $R_\alpha$  and  $R_{max}$  must be upgraded. Retain the newly generated  $M$  cluster. If there is no  $R_\alpha \geq R_{max}$ , then directly cluster the newly generated  $M$ .

(10) The result of this clustering is network behavior awareness.

**3.2. A situation awareness method in a network environment based on NAWL-ILSTM.** An information fusion method based on NAWL-ILSTM is proposed. The extended and short-memory neural networks based on Look-ahead are optimized using the Look-ahead and Naddam algorithms to perceive the network environment [12] effectively. Detailed calculation steps are as follows:

(1) Use  $W = (w_1, w_2, \dots, w_n)$  to represent the accurate timing. The extended  $W$  sequence is a matrix.

$$\begin{bmatrix} w_1 & w_2 & \cdots & w_{x-t+1} \\ w_2 & w_3 & \cdots & w_{x-t+2} \\ & & \vdots & \\ w_t & w_{t+1} & \cdots & w_x \end{bmatrix} \quad (3.2)$$

$x$  is the length of the sequence, and  $t$  is the sample size. In formula (3.3),  $g = (w_t, w_{t+1}, \dots, w_n)$  represents the normalized time series  $W$  of a training sample.

$$W = \frac{w_i}{\sqrt{w_i^2 + w_{i+1}^2 + \cdots + w_{i-t+1}^2}}, i = 1, 2, \dots, x - t + 1 \quad (3.3)$$

(2) Set the network parameters according to formula (3.4), and set the initial value of the network parameters.

$$\begin{cases} V_g = rand(L, F) \\ p_g = rand(1, F) \\ \vdots \\ Max\_iter = Y_1 \\ Error\_Cost = Y_2 \end{cases} \quad (3.4)$$

$V_g$  indicates the forgotten gate weight.  $p_g$  is for forgotten door position deviation.  $F$  is the number of ganglion segments and  $L$  is the number of *LSTM* brain segments.  $Y_1$  stands for the maximum number of iterations  $Max\_iter$ .  $Y_2$  stands for error threshold  $Error\_Cost$ .

(3) The state information  $\widehat{c}_t$  of the unit that needs to be ignored is calculated by formula (3.5).

$$\widehat{c}_t = \mu(V_g * [\zeta_{t-1}, w] + p_g) * \Lambda_{t-1} \quad (3.5)$$

$\mu(V_g * [\zeta_{t-1}, w] + p_g)$  is the output value of the Forget gate.  $\Lambda_{t-1}$  represents the condition of the unit at the last time.

(4) Calculate the amount of information stored in the battery state at A certain time using formula (3.6):

$$\widehat{i}_t = \mu(V_i * [\zeta_{t-1}, w_t] + p_i) * \tan \zeta(V_\Lambda * [\zeta_{t-1}, w_t] + p_\Lambda) \quad (3.6)$$

$\mu(V_i * [\zeta_{t-1}, w_t] + p_i)$  indicates the result of the input port  $i$ . This output determines the value the battery unit needs to be modified.  $\tan \zeta(V_\Lambda * [\zeta_{t-1}, w_t]$  stands for the alternate vector  $\Lambda_t$ , and this alternate vector is created using  $\tan \zeta$ .

(5) The state  $\Lambda_t$  of the battery is obtained from formula (3.7).

$$\Lambda_t = \widehat{i}_t + \widehat{c}_t \quad (3.7)$$

From (7) it is possible to see the state of the cell unit obtained by combining the states of the input grid and the forgotten grid.

(6) The output of the network is calculated using formula (3.8) at t time point:

$$\zeta_t = \mu(V_u * [\zeta_{t-1}, w_t] + p_u) * \tan \zeta(\Lambda_t) \quad (3.8)$$

$\mu(V_u * [\zeta_{t-1}, w_t] + p_u)$  represents the result of the output gate  $u_t$ .  $\zeta_t$  is the confidence of the current point in time.  $\tan \zeta(\Lambda_t)$  represents the condition of the unit. The reliability of the whole training sample is obtained through repeated calculations of (3.3) to (3.6).

(7) The deviation between the total perceived and actual value is obtained from formula (3.9).

$$\Gamma_{(\beta)}(g, \zeta; V, p) = \frac{1}{2} \|g - \zeta\|^2 \quad (3.9)$$

The network is modified by the method of back-to-back propagation [13]. Let's go back to step 3. Up to the maximum number of duplicates or error threshold. If the current number of repetitions  $iter$  is more than  $Max\_iter$ , or  $error$  is more than  $Error\_Cost$ , the training cycle is terminated.

(8) The NAWL algorithm trains the neural network model of ILSTM. Please enter the weighting matrix to be updated  $\delta_0 = [V_c, V_i, V_v, V_u]$ . The initial values are optimized overall when new sampling points are added. This method only needs a simple iteration to get a new optimization result.

(9) Algorithm for dynamic parameter correction based on online observation results [14]. Add new samples  $\delta_0$  and  $W_{n+1}(w_{n-t+2}, \dots, w_{n+1})$ . The perceived value  $\zeta_{n+1}$  of the obtained new sample is propagated forward according to (3.3) (3.6).

$$error = error + \frac{1}{2} (\zeta_{n+1} - w_{n+2})^2 \quad (3.10)$$

(10) Assume that the observed data at the next sampling time reaches where the network was attacked. Relevant network security personnel can detect the alarm the first time the network is attacked and respond quickly [15]. Otherwise, there will be a more profound attack on the network.

**4. Experimental simulation.** The model's empirical analysis verifies the proposed model's effectiveness in practice. Firstly, the network operation data of tobacco Industrial Enterprise in a certain period is collected to form a sample data set for selection. The collected information includes network access information, network speed, and device storage information [16]. The fuzzy degree of the obtained data is analyzed by the fuzzy correlation degree after the data is preprocessed. The results are analyzed by game theory, and the corresponding learning model is established. The topology of the network is shown in Figure 4.1.

The data from the experiment is imported into the learning model, and the data association cluster composed of 30 central nodes is obtained. The test parameters are listed in Table 4.1.

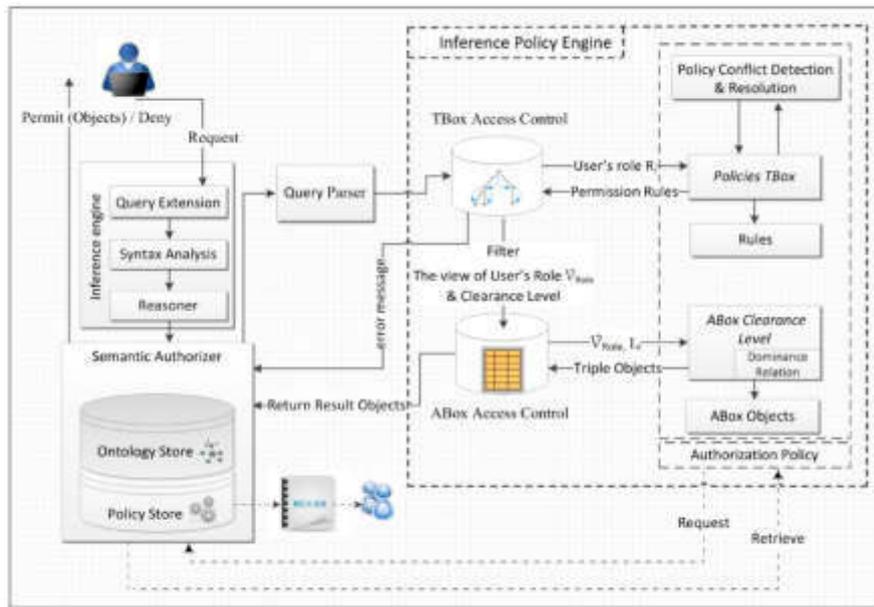


Fig. 4.1: Mesh topology.

Table 4.1: Experimental parameters.

Item	Argument
Computer	Equipped with 2.5GHzIntelI5CPU
Running memory	8GB
Storage hard disk	256GB
Operating system	Windows10

The test data in Table 4.1 can meet the requirements of effective operation. The computer data operation program is used to learn and model network security information to obtain abnormal data information [17]. After eliminating the difference in time, the data's similarity is analyzed by the game method. The original data are compared with the results of relevant analysis. Correct the wrong data operation results, and finally obtain the data information that can reflect the network security situation. The results of the cumulative value at risk are shown in Figure 4.2.

It can be seen from Figure 6 that the deviation of the cumulative risk amount described by the method described in this paper is slight and has little impact on the regular operation of the tobacco Industrial Enterprise. The whole system works pretty steadily, and the network speed is fast. Reliable and complete data were obtained after preliminary preprocessing and preliminary screening [18]. At the same time, the time required is significantly shortened. A risk assessment of the operation of the tobacco Industrial Enterprise during the disposal period was conducted, and the assessment was "low." Experimental results show that the proposed algorithm is efficient and stable. The analysis of this data has little impact on the regular operation of the entire network. In addition, the risk to the network environment is also low. The results of the error rate cognition test are given in Table 4.2.

Through the test of the actual system, it is concluded that the algorithm's error rate proposed in this paper is minimal in practical application. This project will use the NAWL-ILSTM algorithm, game theory, machine learning and other technologies to analyze network security situations. In this system, the data information and similarity factors are operated on many levels, the relationship is established, and the time error problem is reduced to obtain a meager error rate. The results show that the calculation accuracy of the data with multi-

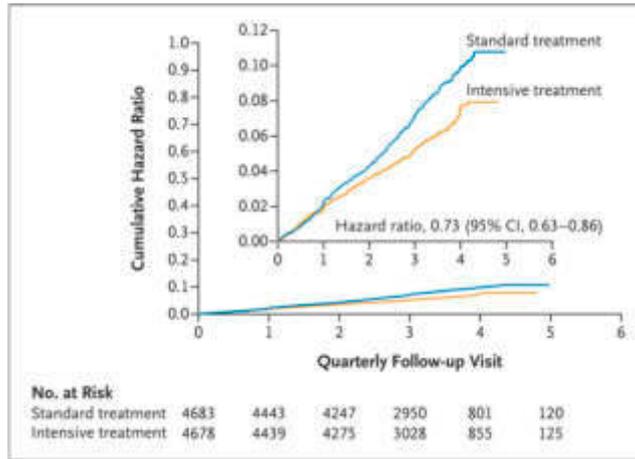


Fig. 4.2: Cumulative var results.

Table 4.2: Error rate perception experiment results.

Number of tests /n	Traditional algorithm	Text algorithm
1	11.71	1.30
2	10.56	1.13
3	12.84	1.27
4	15.88	2.02
5	11.75	1.11
6	13.22	1.70
7	13.79	1.06
8	15.13	1.27
9	14.26	1.42
10	15.50	1.48

layer operation is higher. It can reflect the current network security situation more intuitively and accurately. This project can use various calculation methods to obtain accurate information about network security issues in the environment to ensure the security and stability of the entire network. This method has high efficiency, low risk, and a meager error rate.

**5. Conclusion.** Combining multi-source and heterogeneous data, a network security situation perception system for tobacco Industrial Enterprise in complex environments was established. The system ring’s physical structure and the system layer’s concept model are presented. The real-time monitoring and data mining of tobacco enterprise information systems are realized. The application of data mining technology in the tobacco industry can analyze the security events that may occur in the industry communication and network security platform and the association rules and statistics to realize the comprehensive dynamic monitoring of network security. The method proposed in this project can effectively reduce the system’s running time and improve the system’s reliability for information. The experimental results show that the network security situation awareness technology studied in this project has high practical value in practice.

REFERENCES

[1] Tan, L., Yu, K., Ming, F., Cheng, X., & Srivastava, G. (2021). Secure and resilient artificial intelligence of things: a HoneyNet approach for threat detection and situational awareness. *IEEE Consumer Electronics Magazine*, 11(3), 69-78.

- [2] Zwilling, M., Klien, G., Lesjak, D., Wiechetek, Ł., Cetin, F., & Basim, H. N. (2022). Cyber security awareness, knowledge and behavior: A comparative study. *Journal of Computer Information Systems*, 62(1), 82-97.
- [3] Guo, H., Li, J., Liu, J., Tian, N., & Kato, N. (2021). A survey on space-air-ground-sea integrated network security in 6G. *IEEE Communications Surveys & Tutorials*, 24(1), 53-87.
- [4] Zhang, D., Feng, G., Shi, Y., & Srinivasan, D. (2021). Physical safety and cyber security analysis of multi-agent systems: A survey of recent advances. *IEEE/CAA Journal of Automatica Sinica*, 8(2), 319-333.
- [5] Lv, Z., Qiao, L., Kumar Singh, A., & Wang, Q. (2021). AI-empowered IoT security for smart cities. *ACM Transactions on Internet Technology*, 21(4), 1-21.
- [6] Rosenberg, I., Shabtai, A., Elovici, Y., & Rokach, L. (2021). Adversarial machine learning attacks and defense methods in the cyber security domain. *ACM Computing Surveys (CSUR)*, 54(5), 1-36.
- [7] Bécue, A., Praça, I., & Gama, J. (2021). Artificial intelligence, cyber-threats and Industry 4.0: Challenges and opportunities. *Artificial Intelligence Review*, 54(5), 3849-3886.
- [8] Logeshwaran, J., Ramkumar, M., Kiruthiga, T., & Sharanpravin, R. (2022). The role of integrated structured cabling system (ISCS) for reliable bandwidth optimization in high-speed communication network. *ICTACT Journal on Communication Technology*, 13(01), 2635-2639.
- [9] Zhang, J., Pan, L., Han, Q. L., Chen, C., Wen, S., & Xiang, Y. (2021). Deep learning based attack detection for cyber-physical system cybersecurity: A survey. *IEEE/CAA Journal of Automatica Sinica*, 9(3), 377-391.
- [10] Yang, X., Shu, L., Chen, J., Ferrag, M. A., Wu, J., Nurellari, E., & Huang, K. (2021). A survey on smart agriculture: Development modes, technologies, and security and privacy challenges. *IEEE/CAA Journal of Automatica Sinica*, 8(2), 273-302.
- [11] Butt, O. M., Zulqarnain, M., & Butt, T. M. (2021). Recent advancement in smart grid technology: Future prospects in the electrical power network. *Ain Shams Engineering Journal*, 12(1), 687-695.
- [12] Alizadeh, D., Alesheikh, A. A., & Sharif, M. (2021). Vessel trajectory prediction using historical automatic identification system data. *The Journal of Navigation*, 74(1), 156-174.
- [13] Sadeeq, M. M., Abdulkareem, N. M., Zeebaree, S. R., Ahmed, D. M., Sami, A. S., & Zebari, R. R. (2021). IoT and Cloud computing issues, challenges and opportunities: A review. *Qubahan Academic Journal*, 1(2), 1-7.
- [14] Rawal, B. S., Manogaran, G., & Hamdi, M. (2021). Multi-tier stack of block chain with proxy re-encryption method scheme on the internet of things platform. *ACM Transactions on Internet Technology (TOIT)*, 22(2), 1-20.
- [15] Endsley, M. R. (2021). A systematic review and meta-analysis of direct objective measures of situation awareness: a comparison of SAGAT and SPAM. *Human factors*, 63(1), 124-150.
- [16] Sabireen, H., & Neelananarayanan, V. J. I. E. (2021). A review on fog computing: Architecture, fog with IoT, algorithms and research challenges. *Ict Express*, 7(2), 162-176.
- [17] Georgiadou, A., Mouzakitis, S., & Askounis, D. (2022). Working from home during COVID-19 crisis: a cyber security culture assessment survey. *Security Journal*, 35(2), 486-505.
- [18] Zhou, X., Liang, W., Li, W., Yan, K., Shimizu, S., Kevin, I., & Wang, K. (2021). Hierarchical adversarial attacks against graph-neural-network-based IoT network intrusion detection system. *IEEE Internet of Things Journal*, 9(12), 9310-9319.

*Edited by:* Zhigao Zheng

*Special issue on:* Graph Powered Big Aerospace Data Processing

*Received:* Sep 12, 2023

*Accepted:* Oct 18, 2023



## A DISTRIBUTED SYSTEM FAULT DIAGNOSIS SYSTEM BASED ON MACHINE LEARNING

YIXIAO WANG\*

**Abstract.** Now days distributed system becomes the mainstream system of information storage and processing. Compared with traditional systems, distributed systems are larger and more complex. However, the average probability of failure is higher and the difficulty, complexity of operation and maintenance are greatly increased. Therefore, it is necessary to use efficient methods to diagnose the system. Our aim is to use the trained model to diagnose the fault data of the distributed system, so we can obtain as high diagnostic accuracy as possible, and create a web side for users to use. The technique we proposed uses the integrated learning approach of Stacking to model the superposition of the raw data. To realize this, we trained with a dataset of 10,000 pieces of data and assessed accuracy every once in a while. Our best training results are about 80.69% accurate and can be used on the web side. By training data sets and analyzing distributed system faults with Stacking technology, a model with a test accuracy of 80.69% was obtained. Through this model and the web platform we built, the fault of distributed system can be diagnosed, and the diagnosis results are better than other models.

**Key words:** distributed system; Machine learning; Stacking algorithm; Fault diagnosis; Forecast; Deep learning

**1. Introduction.** Compared with traditional centralized systems, distributed systems have become more and more widely used to deal with large-scale complex data and support various cloud computing platforms [1, 2]. Distributed system is a combination of independent computers, which communicate with each other through the network, share resources, and cooperate with each other to achieve distributed processing [3, 4].

As the distributed system being more widely applied in people's production and life, the problems of distributed system such as large scale, high probability of failure, and difficult to find and diagnose are gradually exposed [5, 6]. Therefore, the key issue is to use technical means to analyze the fault data of the distributed system, design the fault diagnosis model, efficiently analyze and identify the fault categories, then realize the intelligent fault operation and maintenance of the distributed system, quickly recover the fault, greatly reduce the difficulty of the operation and maintenance of the distributed system, and reduce the consumption of human resources [7, 8].

When a node in a distributed system fails, it propagates along the topology, causing the related indicators of adjacent nodes of the distributed system node to fluctuate and a large number of log exceptions. Through these large amounts of fault feature data and label data, these log anomalies can be trained by machine learning, deep learning and other technologies.

The Stacking method that we use consists of two phases: the basic model training phase and the meta model training phase. In the basic model training phase, each basic model is trained using training data. Then, each basic model will predict the verification data and get the predicted result. In the meta model training phase, the predictions of all the underlying models are combined and input into the meta model as new features. This method can deal well with the complex and huge data set in the distributed system fault log, and get more accurate model results.

**2. Model Analysis.** Stacking is an integrated learning method that uses models to model the superposition of raw data. It first learns the raw data through the base learner, the base learner outputs the raw data, and then the outputs of these models are stacked in columns to form new data in the dimensions of  $(m, p)$   $(m, p)$   $(m, p)$   $(m, p)$   $(m, P)$ , and then hand over the new sample data to the second-layer model for fitting.

Figure 2.1 here shows the principle and execution of the Stacking method.

---

\*School of Computer Science and Artificial Intelligence, Wuhan University of Technology, Wuhan, 430070, China (3167201991@qq.com)

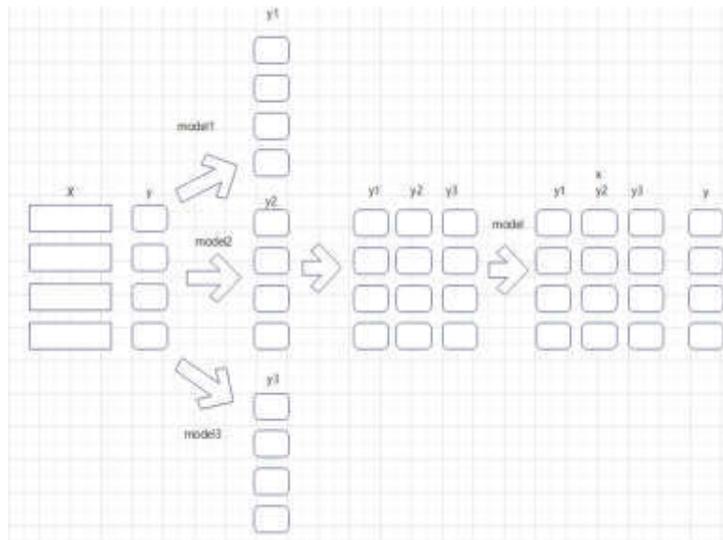


Fig. 2.1: Schematic diagram of Stacking methods.

**2.1. Data Set Introduction.** There are 107 features and labels in this dataset, and there are 6 categories in the labels, which are 0, 1, 2, 3, 4 and 5 respectively. There are 10001 samples in the dataset, but 3704 lines of duplicate data are analyzed and deleted. It can be seen that statistics such as mean value, standard deviation, minimum value and maximum value of data vary greatly among different features.

The correlation coefficients in this dataset range from -1 to 1, where -1 means completely negative correlation, 1 means completely positive correlation, and 0 means no correlation. By analyzing the correlation coefficient matrix, the project can make the following observations:

1. The correlation between the target variable "label" and other features is weak, and the correlation coefficient between most features and the target variable is close to zero.
2. In terms of the correlation between features, there are strong positive correlations between some features. For example, the correlation coefficient between feature0 and feature5 is 0.195971, and the correlation coefficient between feature3 and feature5 is 0.216278.
3. Similarly, there is a strong negative correlation between certain features. For example, the correlation coefficient between feature1 and feature4 is -0.079628, and the correlation coefficient between feature1 and feature5 is 0.021696.

In this dataset, the correlation coefficient between features and labels is screened, but the correlation between most features relatives weakly, with only 3 features higher than 0.2, 7 features higher than 0.1, and 101 features higher than 0.008. After testing, it is unrealistic to select features according to the correlation coefficient. So we need to find another way.

Figure 2.2 shows the data distribution analysis diagram of this dataset.

And Figure 3 here reflects the correlation between features.

A variance of 0 means that all the data are equal. In machine learning, features with zero variance do not contribute anything to the model because they provide no information. Using the zero-variance feature only adds complexity to the model without increasing its predictive power. In this data set, 'feature57', 'feature77', and 'feature100' have variance of 0. Therefore, they need to be deleted.

**2.2. Basic Model And Metamodel.** In this project, multiple classifiers are used for model fusion, among which the three basic classifiers are CatBoost, random forest and support vector machine, and the meta-classifiers are logistic regression. These classifiers are grouped together and the stacking method is used for model fusion.

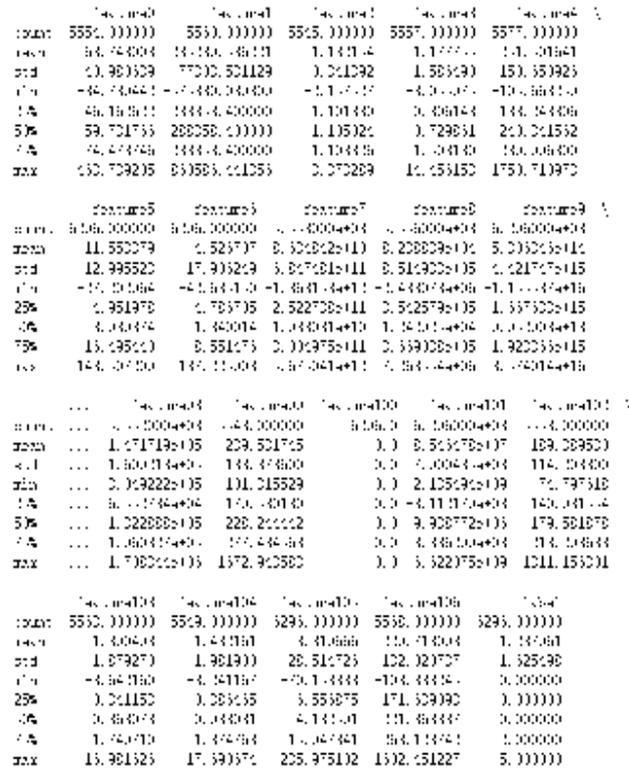


Fig. 2.2: Data distribution analysis diagram.

Here are brief introductions to the base learner and metamodel:

1. CatBoost is a gradient lifting decision tree algorithm developed by Yandex. It is an ensemble learning algorithm capable of handling classification and regression problems. CatBoost builds the model by iteratively training the decision tree, with each iteration reducing the prediction error from the previous iteration. CatBoost has many advantages, for example: (1) Robustness: CatBoost can handle different types of data, including numeric, categorical, and textual data. It can also handle missing values and outliers; (2) Efficiency: CatBoost uses a number of optimization techniques to accelerate model training and prediction, including symmetric trees, semi-sorting and fast histogram algorithms; (3) Accuracy: CatBoost uses techniques such as order lifting and category feature combination to improve the accuracy of the model; (4) Easy to use: CatBoost provides rich API and documentations that supports multiple programming languages and platforms. It also provides a number of pre-processing tools and visualization tools for ease of use. This project uses the CatBoostClassifier class to create a CatBoost classifier that uses Gpus for training and multi-classification evaluation metrics with 200 iterations, a learning rate of 0.09, and an early stop technique to prevent overfitting.

2. Random forest is an integrated learning algorithm based on decision trees, which can deal with classification and regression problems. Random forest builds models by constructing multiple decision trees, each of which will predict data, and the final prediction result will be decided by voting of the prediction results of all decision trees. Random forest has many advantages, including: (1) robustness: Random forest can handle different types of data, including numerical and categorical data, it can also handle missing values and outliers; (2) Accuracy: Random Forest improves the accuracy of the model by constructing multiple decision trees, and it also uses techniques such as self-help method and feature sampling to reduce overfitting and underfitting problems; (3) Interpretability: Random forests can provide importance scores to help users understand the contribution of each feature to the predicted outcome. This project uses the RandomForestClassifier class in the scikit-learn library to create a random forest classifier. The classifier is trained using default parameters.

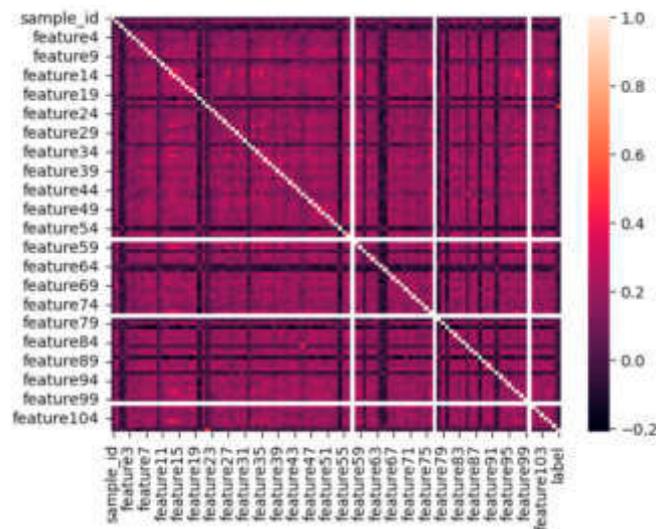


Fig. 2.3: Correlation Coefficient Heat Map.

3. Support Vector Machine (SVM) is a supervised learning algorithm that can handle classification and regression problems. SVM maximizes the space between different classes by finding a hyperplane to separate them. Support vector machines can also use kernel functions to map data into high-dimensional Spaces to solve nonlinear classification problems. This project uses the SVC class from the scikit-learn library to create a support vector machine classifier. This classifier uses radial basis kernel functions and turns on probability estimation.

4. Logistic regression is a generalized linear model that can handle binary classification problems. In the code of this project, the logistic regression model is used as a meta-classifier to combine the predictions of the underlying classifier.

5. The Stacking method is an integrated learning method that combines the prediction results of multiple basic models to improve the accuracy of the model. The Stacking method consists of two phases: the basic model training phase and the metamodel training phase: In the basic model training phase, each basic model is trained by using training data. Then, each basic model will predict the verification data and get the predicted result. In the metamodel training phase, the predictions of all the underlying models are combined and input into the metamodel as new features. The metamodel is trained using these new features and the real labels of the validation data, and finally, the metamodel makes predictions on the test data to get the final predictions. In this project, CatBoost classifiers, random forest classifiers, and support vector machine classifiers are used as the base classifiers, and logistic regression models are used as meta-classifiers. These classifiers are grouped together and the model fusion is performed using the stacking method.

**2.3. Analysis of Stacking Model.** First, feature  $x$  and label  $y$  are put into the three models, and then each model is learned separately. The value of  $x$  is then predicted, sometimes given the proba probability, where we use the predicted value and then overlay the output values of the three models to form a new sample data. The new sample data is then used as label  $x$ , the label of the new data is still the label  $y$  of the original data, and the  $x$  and  $y$  of the new data are passed to the second layer model for fitting, which is used to fuse the results of the previous round of three models.

However, this model often overfits, so the above method is improved by using K-fold cross-validation method. The difference is that each model in the figure is trained on all the data and then outputs  $y$  to form new data. K-fold cross-validation is used to train only K-1 folds at a time, and then the remaining predicted value of 1 fold is used as new data.

K-fold cross-validation is used to divide the data into 4 folds to form 4 sets of data sets, with yellow

representing the training set and green representing the verification set. Then, each set of training sets is given to the model for training, and the verification set is predicted to obtain the output of the corresponding verification set. Since the data is divided into 4 groups by 4 times of cross-validation, there will be 4 verification sets. The predictions for each model are then stacked on their own validation set, row by row, to get the predicted values for the full sample data. Each model takes the predicted values in this way and then merges them into the columns.

In this project, for a single model, class 0 and Class 1 are easily confused, and the classification effect is not significant, or the classification effect of class 4 is not good, such as catboost in class 4 classification effect is good, but class 0 and class 1 classifications are not good, support vector machine in class 0 classification effect is good, but in class 4 classification is not ideal. Therefore, the fusion method of stacking models is used to integrate three basic learning tools, catboost, random forest, and support vector machine, to play their respective advantages and improve the overall accuracy.

**3. Web Design.** The system is divided into web side and server side, using their interaction to complete the data collection, processing and result return. The web side can obtain the file data from the user and transfer the file to the server side. The server receives the file, performs various operations, then returns the result.

**3.1. Front Design.** The front-end part mainly uses html, css, JavaScript design and jQuery to improve efficiency, and the design results can be stably displayed on the web.

We have carried on the function analysis and determined the various front-end interfaces required for this function. It can be divided into login interface, function realization interface and information introduction interface.

Among them, the login interface is divided into registration and login two functions, and use the button to switch. The registration function requires the user to enter the user name, password and other information, the web page can determine whether the format of the user input is correct and prompt, when the user input correctly and confirm the registration, the data is passed into the background and stored in the database. Login function requires the user to enter the corresponding user name and password, the web page can determine whether the input format is correct and prompt, and then the pair of user name and password into the back-end to determine whether the corresponding. If the username and password are correct, the user can enter the function implementation interface.

The functional implementation interface is divided into two parts: training data and test samples. Training data allows users to upload a.CSV format file, and then the file will be sent to the background for training, and directly download the model file after successful training. The Test sample section allows the user to select a.joblib format model file and a.csv file for testing, click Start Training, and then use the model to train the dataset in the background. After the training, click the test result preview button to jump to the result preview interface to view the training results.

The information introduction interface includes the distributed system introduction interface, and the Personal information interface. Distributed Systems Introduction interface provides basic information about distributed systems and their diagnostics. The Personal Information interface stores information of the user who has logged in and allows the user to modify his or her information.

**3.2. Back Design.** The back-end part uses the mainstream python framework django, which has been popular for a long time. It has many functions, for example, the system is complete, which can help complete the design quickly and development of web pages.

First, using django, you can generate the files directly from the command line by using the startproject to create the main project and startapp to create the user application.

In the second step, you need to do this in settings.py in the main project. First add the application and middleware that will be used, then set the database interface to mysql and select the user and password to log in to mysql and the database to be used. In this case, select a new user database as the database of the web application. Then set the static file path so that you can better call css, js, and image ICONS from the templates html file. Finally, a compress application interface is set up to better check the effect of css and js

```

validate_1000.csv
model_id_str.joblib
test_f1 score: 0.886977837179146
F1 score for each class: [0.69883546 0.63629515 0.64122137 0.96969697 0.9382716 0.96153846]

```

	precision	recall	f1-score	support
0	0.60	0.84	0.70	176
1	0.61	0.67	0.64	166
2	0.92	0.49	0.64	171
3	0.99	0.95	0.97	169
4	0.98	0.97	0.94	136
5	1.00	0.93	0.96	162
accuracy			0.88	1000
macro avg	0.84	0.81	0.81	1000
weighted avg	0.83	0.80	0.80	1000

Fig. 4.1: Result Model.

changes in development, so that every change in css and js can be displayed instead of using the system cache content.

The third step is to call the urls in the user file in the urls.py setting of the main file, and set the urls to be used in the user file and the function and name that it calls in views.py.

The fourth step is to declare the properties of the user in the models.py and apps.py files of the user file, and use migrate to allow the system to create the corresponding tables directly in the user database, which also makes it easy to access the database in the subsequent view functions.

The fifth step is to define the function used in views.py, the view function. First, form is declared to process the user information input by the front end. Other information, such as the file information uploaded by the user, can be obtained through request. The view function is mainly written in the login function and the main interface of the index function. The login function will compare the user's information with that of the user in the database in the registration section, so that the registration can be carried out without having repeated user name and email address. The registered user information will also be stored in the database, and the user name or email address entered by the user can be queried in the login section. Compare the password stored in the database with the password entered by the user. If the password is the same, the user can log in successfully and go to the main page. In the index function, the training part and the test part are mainly processed. In the training part, the file selected by the user is obtained and the my\_train function is called for training, and the training model file is downloaded at last. In the test part, the test machine file and model file selected by the front-end user are obtained and the test function is called for testing. Finally return to download the test result file.

**4. Results and Discussion.** This project studied how to use machine learning to diagnose faults in distributed systems. We used appropriate Stacking algorithms, trained sample fault logs, and improved them constantly to obtain a relatively accurate result model with an accuracy rate of more than 80% on the verification set.

The result analysis of the validation set and its accuracy rate are shown in Figure 4.1.

Also, the system can run on the web platform, allow users to upload training data and train online, and download the obtained model after completion. It also allows single or batch upload of test samples, and the results can be visualized and downloaded.

#### REFERENCES

- [1] Chieh-feng Chiang, Tan, J.J.M. Using Node Diagnosability to Determine t-Diagnosability under the Comparison Diagnosis Model [J]. IEEE Transactions on Computers, 2009, 58(2): 251-259
- [2] F Sun, E Xu, H Ma. Design and comparison of minimal symmetric model-subset for maneuvering target tracking [J]. Journal of Systems Engineering and Electronics, 2010, 21(2): 268-272.

- [3] Zhi-Hua Zhou, Yuan Jiang, Medical diagnosis with C45 Rule preceded by artificial neural network ensemble [J]. IEEE Transactions on Information Technology in Biomedicine, 2003, 7(1):37-42.
- [4] Fronza I, Sillitti A, Succi G, et al. Failure prediction based on log files using random indexing and support vector machines [J]. Journal of Systems & Software, 2013, 86(1):2-11.
- [5] Lan Z, Gu J, Zheng Z, et al. A study of dynamic meta-learning for failure prediction in large-scale systems [J]. Journal of Parallel & Distributed Computing, 2010, 70(6):630-643.
- [6] Mi HB, Wang HM, Zhou YF, et al. Localizing root causes of performance anomalies in cloud computing systems by analyzing request trace logs [J]. Science China Information Sciences, 2012, 55(12):2757-2773.
- [7] Rao X, Wang HM, Chen ZB, et al. Detecting faults by tracing companion states in cloud computing systems [J]. Chinese Journal of Computers, 2012, 35(5):856-870.
- [8] Yu X, Joshi P, Xu J, et al. CloudSeer: Workflow monitoring of cloud infrastructures via interleaved logs [J]. SIGOPS Operating Systems Review, 2016, 50(2):489-502.

*Edited by:* Zhigao Zheng

*Special issue on:* Graph Powered Big Aerospace Data Processing

*Received:* Sep 25, 2023

*Accepted:* Oct 11, 2023



## INTELLIGENT NAVIGATION SYSTEM BASED ON BIG DATA TRAFFIC SYSTEM

XIU ZHANG\*, JIAN KANG<sup>†</sup> AND HAICUN YU<sup>‡</sup>

**Abstract.** This paper studies the navigation technology of the On-Orbit Servicing Spacecraft (OSS) and proposes an overall solution for the OSS navigation system. The composition of the OSS satellite navigation simulation system and simulation supporting environment are studied. Then a new SSUKF navigation filter is proposed based on the hyper spherical distributed feature sampling point transform algorithm (SSUT) and the non-scanning Kalman filter (UKF). The numerical simulation of the OSS system is studied based on MATLAB/RTW. Finally, the SSUKF algorithm and the conventional UKF algorithm are simulated digitally. The effectiveness and advanced nature of the hybrid Kalman filter applied to spacecraft autonomous navigation are verified.

**Key words:** Spacecraft; Big data; Hyperspherical distributed sampling point transformation algorithm; Unscented Kalman filtering; Navigation system

**1. Introduction.** A spacecraft's autonomous navigation system uses its own load or observation instrument to realize the precise positioning of the spacecraft and motion trajectory. Since 1963, the United States has carried out in-depth research on the autonomous navigation of satellites and developed various types of sensitive elements and space navigation systems. In recent years, significant progress has been made in studying small satellites and microsatellites in China. At present, autonomous navigation has gradually become one of the leading indicators to measure the comprehensive capability of spacecraft. It is found that the GNC system has the highest failure probability based on the statistics of spacecraft failure probability data in the recent ten years. A navigation system is an essential part of the GNC system. The "rendezvous" and "docking" part of the open satellite is the crucial part of the whole flight satellite navigation system [1]. The approach segment of the image guidance scheme employed by OE uses infrared cameras (IRS) to back up advanced image guidance sensors (AVGS). But the IRS's adequate scope does not cover all approximate segments. When AVGS fails during the approach process, replacing the spare sensitive element is unnecessary to perform the on-orbit mission. The vehicle uses RGPS navigation. Communication between spacecraft is essential. Due to the large number of non-cooperative targets in space, coupled with the existing domestic relay communication system has not reached, the world can only use RGPS technology.

Unsamped Kalman filter (UKF) is a state estimation method based on unsampled transformation. The nonlinear control method based on UKF is superior to the traditional Kalman filter and has a broad application prospect. The calculation speed of the UKF method mainly depends on the number of sample points in the unsampled transformation. The conventional UKF algorithm performs unsampled conversion on  $2n+1$  sampling points. The sampling point is symmetric concerning the mean of the  $n$ -dimensional vector. Because the vector dimension is too large, the operation speed of the algorithm will be reduced [2]. This project will adopt Superball Distribution Transform (SSUT) to replace the traditional symmetric distribution. The algorithm can effectively reduce the number of samples and improve the algorithm's efficiency without affecting the filtering effect. The thesis uses MATLAB/RTW as the platform to study the positioning of OSS. A new SSUKF navigation filter is proposed by fusing SSUT and UKF. The algorithm is applied to the autonomous navigation algorithm of spacecraft using starlight Angle distance as observation information.

## 2. The overall scheme of the OSS navigation simulation system.

---

\*Department of Information Technology, Tangshan Vocational and Technical College, Tangshan, 063000, China

<sup>†</sup>Department of Information Technology, Tangshan Vocational and Technical College, Tangshan, 063000, China

<sup>‡</sup>Department of Information Technology, Tangshan Vocational and Technical College, Tangshan, 063000, China (Corresponding author, [tsvtczhangxiu@126.com](mailto:tsvtczhangxiu@126.com))

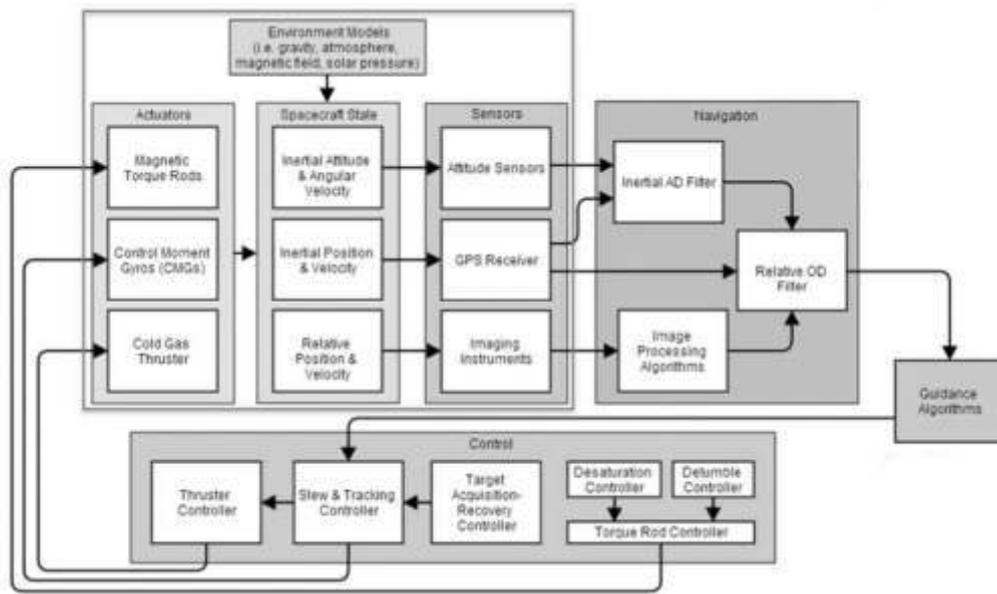


Fig. 2.1: Overall simulation architecture of OSS navigation system.

OSS Navigation System Mission Planning	Simulation support environment based on MATLAB/RTW
Tracking Spacecraft Orbit Mission Planning	Simulation run process control
Target spacecraft orbit mission planning	Simulation running status monitoring
Spacecraft on-orbit mission planning	Simulation Data Management
Spacecraft Rendezvous Mission Planning	Communication settings and management between data
Tracking Spacecraft Propulsion Strategy Planning	Simulation Data Visualization
Strategy planning for autonomous operation of navigation system in orbit	
OSS Navigation System Simulation	OSS Navigation System Performance Evaluation
OSS Orbital Dynamics	System Integrity Performance Assessment
Target spacecraft orbital dynamics	GNSS Constellation Performance Evaluation
Satellite Navigation Simulation	Performance Evaluation of Spaceborne GNSS Receiver
Inertial Navigation Simulation	Performance Evaluation of Spaceborne Inertial Measurement Units
Celestial Navigation Simulation	Performance Evaluation of Spaceborne Celestial Navigation Measurement Components
Data Processing and Information Fusion	Performance Evaluation of Spaceborne Relative Navigation Measurement Components
Relative Navigation Simulation	Performance Evaluation of Spaceborne Image Navigation Measurement Components
Image Navigation Simulation	Performance Evaluation of Spaceborne GNSS Navigation
Space Environment Effects and Observational Data Generation	Performance Evaluation of Spaceborne Inertial Navigation
Navigation data runs autonomously	Performance Evaluation of Spaceborne Celestial Navigation
	Performance Evaluation of Spaceborne Relative Navigation
	Performance Evaluation of Spaceborne Image Navigation
	Navigation Mode and Integrated Navigation Performance Evaluation

Fig. 2.2: Functional structure.

**2.1. Simulation system architecture.** Using MATLAB/RTW as the platform, the simulation of the OSS in-orbit mission, navigation system requirements, navigation sensor setup and navigation mode is carried out. Its overall simulation architecture is shown in Figure 2.1. The functional structure of the OSS navigation simulation system is shown in Figure 2.2. The simulation system mainly includes the following four parts:

**2.1.1. OSS Navigation System Task Planning.** A scheme for tracking the flight path between the spacecraft and the target spacecraft is proposed. The planning problems of spacecraft in orbit and rendezvous flight are studied [3]. It lays a foundation for tracking the propulsion and autonomous operation of the satellite in orbit and its motion trajectory. The satellite's flight path will also provide a necessary reference for the regular operation and test of the simulator.

**2.1.2. Simulation of OSS navigation system.** The OSS navigation system is simulated according to the flight plan. The comprehensive performance of the system under various navigation modes, error influencing factors and coordination of subsystems are studied [4]. The OSS navigation system consists of 10 sub-modules: OSS orbit dynamics, target spacecraft orbit Dynamics, satellite navigation, inertial navigation, astronomical navigation, data processing and information fusion, relative navigation, image navigation, space environment impact and observation data generation, and navigation data autonomous operation.

**2.1.3. Support simulation system implemented by MATLAB/RTW.** The supporting environment provides an operational platform for the system's planning, simulation and evaluation. According to the requirement of a real-time system, the system is divided into two categories: real-time class and non-real-time class [5]. The supporting environment can configure the system in real-time for various real-time requirements. Some basic algorithms, noise simulation and data processing are provided.

**2.1.4. Performance evaluation of OSS navigation system.** The performance of the system is evaluated based on numerical simulation. The performance evaluation of OSS based navigation sensor and navigation system is completed. Various fault and error factors are analyzed [6]. It lays a theoretical foundation for the research of semi-physical and physical simulation.

## 2.2. Key system technologies.

**2.2.1. Synchronizing Data with Time Series.** Each sensor has its own time. Taking GNSS as an example, the target's pose is obtained by the observed pseudo-distance/pseudo-distance rate combined with the star vector observed by STR. Their calculations are based on their clocks. This paper designs a unique clock for the simulation system. Each subsystem has orders. Turn functionality into your clock. It is necessary to divide the timing of each module reasonably to ensure that the simulation performance of each sensor and filter element in the OSS system simulation meets specific causes and consequences. The causal relationship between data transmission synchronization and data transmission needs to be clarified. Synchronous transmission is after OSS generates new trajectory information [7]. The OSS generates new trajectory information, and the receiver generates new trajectory information—the message flow limits data transmission synchronization. It is necessary to use the orbit data of OSS. This paper presents a guidance algorithm based on fuzzy logic to ensure the correctness of the information transmission.

**2.2.2. Fault Detection, Isolation, and Repair (FDIR).** Open-source analysis software is a kind of nonlinear system with dynamic properties. The operation of sensors and actuators must be monitored in real-time during simulated flight. FDIR guarantees the autonomy of an open service system. The system can monitor and analyze the operation of open-source systems [8]. Evaluate the current sensor performance and feedback OSS performance to MVM. An intelligent positioning technology based on OSS is proposed and verified by experiments. An intelligent diagnosis method based on fuzzy logic is proposed. The model uses database technology to express the required knowledge in a particular form. Thus, a knowledge base of expert systems is formed [9]. An expert system based on logical reasoning is constructed using a particular programming language. The knowledge database is maintained and updated.

**2.2.3. Guide mode conversion method.** The primary purpose of navigation mode switching is to ensure the precise and safe switching of OSS in orbit. The navigation system model is managed based on events, moments, and observations from navigation sensors in the OSS flight plan [10]. Determine the current health of open-source software. Triggers the corresponding guidance sensor to start and transmit pattern initialization parameters. Use related data as separate states to express different browsing patterns. Events are set as status start and transfer operations. Develop processes for state transitions in event-driven situations [11]. Transform between navigation modes with State flow and embed it in Simulink. Convert it into an executable program via RTW and download it to the destination system.

## 3. SSUKF navigation filtering algorithm.

**3.1. Superball distribution sampling conversion UKF.** The navigation filtering system of generalized Kalman filtering technology is widely used. However, the generalized Kalman filter has problems, such as difficult debugging, unstable performance when it does not meet the linearization conditions, and the need to

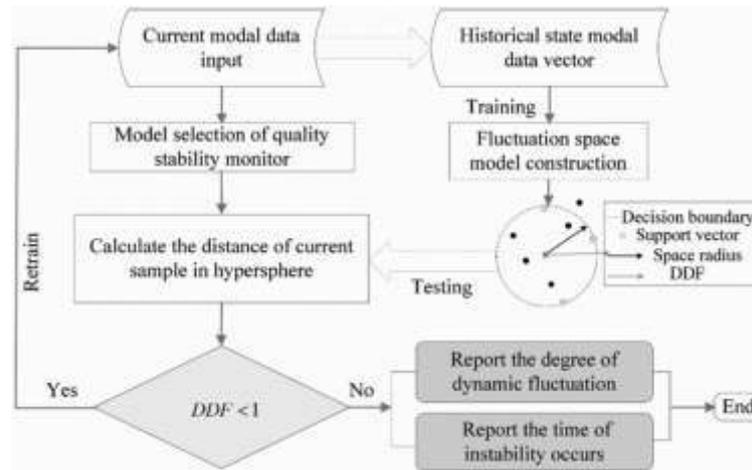


Fig. 3.1: Flow of hyper spherical distributed sampling transformation algorithm.

calculate the Jacobian matrix. The University of Oxford introduced the Superball Point Force Family (UKF) method in this field in the 1990s to overcome the above defects. UKF uses a set of recursive equations similar to Kalman filters. The prior knowledge is used to iterate the system state and system parameters. Then the system parameters are modified according to the relevant knowledge of the observation time point to achieve the average of the system state and parameter estimation. UKF differs from an extended Kalman filter in using a set of sample points. These sample points are close to the Gaussian distribution [12]. The method uses a non-minimization transformation to recurse and modify the covariance of the system state and deviation, thus avoiding the solution of the Jacobian matrix of the system state and the observation system. It is also relatively easy to implement. The basic idea of the EKF algorithm is to carry out Taylor expansion of the system at the measured point without considering the higher order terms of the second order or more [13]. The UKF algorithm approximates the state of the system through non-segmentation transformation. Compared with the EKF algorithm, this algorithm can achieve quadratic or even higher accuracy. Therefore, the accuracy of this algorithm is higher than that of the EKF algorithm. In addition, UKF also adds unconstrained transformation in iteration so that the adjustment coefficient required in an iteration is more than EKF. Its debugging is more flexible. The UKF method is similar to the EKF method in that both are estimates of serialized state parameters [14]. The recursive steps of the structure flow are also composed of the propagation of state and covariance and the updating of measurements. Figure 3.1 shows the Flow of the hyper spherical distributed sampling transformation algorithm (The picture is quoted in the Journal of Process Mechanical Engineering, 2019; 233(3):436-447).

**3.2. SSUT.** Suppose  $x$  is a  $n$  dimensional random variable. The mean matrix is  $\hat{x}$  and the mean square error matrix is  $Q_{xx}$ .  $x$  random variable  $y$  of  $y = g(x)$  dimension is obtained after a nonlinear transformation  $y = g(x)$  is performed on  $x$ . The non-sampling conversion method is to select  $x$  set of samples that can represent the mean and variance of sample  $x$ . By using this method, a nonlinear transformation of  $y$  is carried out, and the converted sample is used to approximate  $y$ .

For non-scan conversion,  $2n + 1$  sample points with uniform distribution are generally selected [15]. It is generally required to have more than  $n + 1$  samples to describe the mean and variance number of  $n$ -dimensional state variables. The probability distribution of the system is approximated by taking  $n + 1$  samples with the same weight on the hypersphere. This sample point is combined with the mean of its states to form a  $n + 2$  sample point without sampling transformation. The mean square error matrix of  $n$  dimension  $z$  is an identity matrix, and the sampling method on the hypersphere is as follows:

1) Option value  $\omega_0$ .

$$0 \leq \omega_0 \leq 1 \tag{3.1}$$

2) Determine weight  $\omega_i$ :

$$\begin{aligned} \omega_i &= (1 - \omega_0)/(n + 1) \\ i &= 1, \dots, n + 1 \end{aligned} \tag{3.2}$$

3) Initialize vector sequence:

$$\left. \begin{aligned} R_0^1 &= [0] \\ R_1^1 &= [-1/\sqrt{2\omega_1}] \\ R_2^1 &= [1/\sqrt{2\omega_1}] \end{aligned} \right\} \tag{3.3}$$

4) Extension vector sequence ( $j = 2, \dots, n$ ):

$$R_i^j = \begin{cases} \begin{bmatrix} R_0^{j-1} \\ 0 \end{bmatrix} & i = 0 \\ \begin{bmatrix} R_i^{j-1} \\ -1/\sqrt{j(j+1)\omega_j} \end{bmatrix} & i = 1, \dots, j \\ \begin{bmatrix} 0^{j-1} \\ j/\sqrt{j(j+1)\omega_j} \end{bmatrix} & i = j + 1 \end{cases} \tag{3.4}$$

$R_i^j$  represents the  $i$  sampling point of a  $j$  dimensional random variable.  $0^j$  stands for  $j$  dimensional zero vector.  $n + 2$  sampling points  $R_i^n, i = 0, \dots, n + 1$  of  $z$  are obtained by the above method [16]. The number of super-spherical distribution samples for a  $n$  dimensional random variable  $x$  with mean value of  $\hat{x}$  and mean square error matrix of  $Q_{xx}$  can be obtained from formula (3.5):

$$X_i^n = \hat{x} + \sqrt{Q_{xx}} R_i^n \quad i = 0, \dots, n + 1 \tag{3.5}$$

**3.3. UKF algorithm based on SSUT.** Suppose that the equations of state and measurement for a nonlinear system are expressed in the following discrete form:

$$x_{t+1} = G(x_t, A_t, t) + \delta_t \tag{3.6}$$

$$y_t = H(x_t, t) + \varepsilon_t \tag{3.7}$$

$x_t$  is the system state vector.  $A_t$  is the input control vector.  $y_t$  is the observation vector.  $\delta_t$  and  $\varepsilon_t$  are the system noise vector and the measurement noise vector respectively. The mean is 0. The variance matrix is denoted by  $Q_t$  and  $S_t$ , respectively. Let the system state vector dimension be  $n$ .

**3.4. Mathematical model of the autonomous navigation system.**

**3.4.1. Spacecraft orbital motion model.** The geocentric equatorial inertial coordinate system establishes the spacecraft orbital motion equation. A two-body motion model with harmonic terms, including the second-order perturbation of the earth’s gravitational field, is taken [17]. The other perturbation factors are

equivalent to Gaussian white noise.

$$\left. \begin{aligned} \frac{dx}{dt} &= \varepsilon_x \\ \frac{dy}{dt} &= \varepsilon_y \\ \frac{dz}{dt} &= \varepsilon_z \\ \frac{d\varepsilon_x}{dt} &= -\lambda \frac{x}{s^3} \left[ 1 - H_2 \left( \frac{S_R}{s} \right)^2 \cdot \left( 7.5 \frac{z^2}{s^2} - 1.5 \right) \right] + \delta_{\varepsilon_x} \\ \frac{d\varepsilon_y}{dt} &= -\lambda \frac{y}{s^3} \left[ 1 - H_2 \left( \frac{S_R}{s} \right)^2 \cdot \left( 7.5 \frac{z^2}{s^2} - 1.5 \right) \right] + \delta_{\varepsilon_y} \\ \frac{d\varepsilon_z}{dt} &= -\lambda \frac{z}{s^3} \left[ 1 - H_2 \left( \frac{S_R}{s} \right)^2 \cdot \left( 7.5 \frac{z^2}{s^2} - 4.5 \right) \right] + \delta_{\varepsilon_z} \end{aligned} \right\} \quad (3.8)$$

$s = \sqrt{x^2 + y^2 + z^2}$ ;  $x, y, z$  is the position coordinate of the spacecraft.  $\varepsilon_x, \varepsilon_y, \varepsilon_z$  is the velocity coordinate of the spacecraft.  $\lambda$  is the gravitational constant of the earth.  $H_2$  is the second order harmonic term coefficient of the earth.  $S_R$  is the earth's equatorial radius.  $\delta_{\varepsilon_x}, \delta_{\varepsilon_y}, \delta_{\varepsilon_z}$  is system noise.

**3.4.2. Starlight angular distance measurement model.** The angular distance of starlight can be determined according to the measurement information of the infrared earth sensor and star sensor. The angle between the vector direction of the light of the navigation star and the vector direction of the earth's center in the spacecraft body coordinate system. Starlight angular distance remains constant in different coordinate systems. The angular distance measured in the body coordinate system can also be used in the equatorial inertial coordinate system [18]. In the geocentric equatorial inertial coordinate system, the expression of starlight angular distance between the two navigation stars is

$$\theta_1 = \beta_1 + \varepsilon_{\theta_1} = \arccos \left( -\frac{x B_{x_1} + y B_{y_1} + z B_{z_1}}{s} \right) + \varepsilon_{\theta_1} \quad (3.9)$$

$$\theta_2 = \beta_2 + \varepsilon_{\theta_2} = \arccos \left( -\frac{x B_{x_2} + y B_{y_2} + z B_{z_2}}{s} \right) + \varepsilon_{\theta_2} \quad (3.10)$$

$B_{x_1}, B_{y_1}, B_{z_1}$  and  $B_{x_2}, B_{y_2}, B_{z_2}$  are the direction cosine of the two navigation stars in the geocentric equatorial inertial coordinate system, respectively.

**4. SSUKF-EKF hybrid filter design.** The UKF method does not need to linearize the observation equation and the equation of state when solving nonlinear problems. The method has no truncation error of higher-order terms. The algorithm has fast convergence and high steady-state accuracy. At the same time, the method has strong robustness to random disturbance. UKF algorithm needs to process multiple sampling points simultaneously in a period, which leads to slow algorithm convergence. EKF algorithm is an efficient algorithm. At the same time, the method has higher accuracy in the case of weakening the nonlinearity [19]. However, the EKF will diverge when there are large initial values and random noise. Since the operation steps of the two methods are based on the Kalman filter, the two methods can be combined to construct the Kalman filter. The Flow of the SSUKF-EKF hybrid filter algorithm is illustrated in Figure 4.1 (image quoted in Journal of neural engineering.13.10.1109 / ICVR.2015.7358598.)  $M$  is a function that increases as the accuracy of the parameter estimation increases. This paper presents the EKF method. If the value of a function exceeds an error threshold, it indicates that the filtering effect of the function becomes worse to improve the computing speed. An improved SSUKF method is proposed. Kalman filter has strong robustness in practical applications, and its stability depends on the state of the system and the measurement equation. Its stability is good if the Kalman filter has random controllability and random observability. The method uses the modal conversion

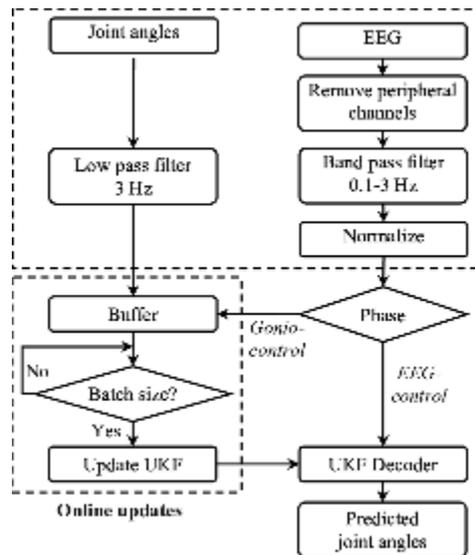


Fig. 4.1: Hybrid Kalman filter algorithm flow.

function and error threshold of EKF and UKF to realize the adaptive conversion of the two algorithms. This is only a few changes to the mean and variance matrix operation and does not affect its stability [20]. Therefore, selecting the proper minimum root mean square can improve the filtering effect while ensuring accuracy.

## 5. System implementation and numerical simulation.

**5.1. System Implementation.** This paper studies the OSS navigation simulation system based on MATLAB/RTW. RTW is an auxiliary graphical model and simulation system. The software converts the Simulink model into an improved, portable, personalized program [21]. Different application scenarios can be generated automatically based on different object structures. Real-time hyper spherical communication technology communicates with multiple targets. RTW provides an accelerated simulation method for processing batch simulations, such as the Monte Carlo shooting type. In addition, the external mode Simulink of RTW can communicate with the model system in the real-time simulation environment to achieve real-time simulation control, operating state monitoring, parameter adjustment and data observation. The architecture block diagram for rapid prototyping of the OSS navigation system using MATLAB/RTW is shown in Figure 5.1 (Electronics 2022,113,1979).

This paper uses MATLAB and Simulink toolkits to study the navigation simulator's overall scheme. The models of each subsystem are established. The flow chart of the system is drawn, and numerical simulation is carried out. The Monte Carlo statistical method is used to evaluate and analyze the performance [22]. The proposed method is improved and applied to the navigation system. The application of Simulink in practical applications is realized using RTW technology—real-time functional verification in external mode.

**5.2. Numerical Simulation.** The equatorial inertial system at the center of the H2000 Earth is used as coordinates. The parameters of the initial orbit are as follows: Long semi-axis  $a=7151.918$  km—eccentricity  $e=0.5$ . The earth's tilt is  $98.52^\circ$ . The correct longitude of the ascending node is  $234.083^\circ$ . The perigee oscillation Angle is  $234.083^\circ$ . Horizontal approach points  $240.817^\circ$ . Sensitivity accuracy includes 1 "star-containing and 0.01-degree infrared Earth sensitivity accuracy. The measurement noise is treated as a normal distribution. The aspherical gravity perturbations of the earth are calculated in the actual operation [23]. In this paper, the earth gravity model is established by using JGM3. The fourth-order Rungkat method is proposed. The numerical simulation of the problem is carried out in the case of solar and lunar gravity, atmospheric drag, solar light pressure and other disturbances.

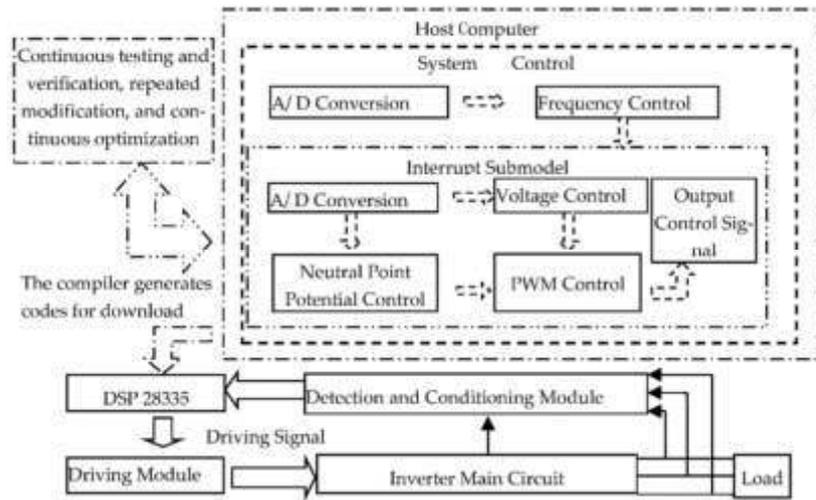


Fig. 5.1: A rapid prototyping method based on RTW.

Table 5.1: Guidance simulation with slight initial deviation and small step length.

Filtering method	Standard UKF	SSUKF	EKF
Position error /m	253.729	253.531	271.229
Velocity error /(m · s-1)	0.267	0.267	0.281
Convergence time /s	2466	2464	2469
Relative computation time	1.042	0.719	0.25

Suppose  $\Delta X_0 = [2 \ -2 \ 2 \ -2 \ 2 \ -2]$  the state variable has an initial deviation of a. The positioning error is defined in km. The error of the speed is expressed as  $m \cdot s^{-1}$ . The simulation takes 15 seconds [24]. The simulation time is four tracks. The simulation results are shown in Table 5.1 and in Figures 5.2 and 5.3. The deviation from the positioning to the rate is defined as the modulus of the deviation vector. The convergence time is the interval between the starting point and the critical point at which the estimated accuracy reaches 1 km. The results show that the SSUKF and UKF algorithms have good filter design consistency. The calculation speed of the algorithm is greatly improved. The EKF method is also the most effective and accurate in the solution process.

**6. Conclusion.** The work and operation process of OSS in orbit is comprehensively analyzed. The structure and working mode of the navigation system is given by selecting the OSS positioning sensor. A novel Kalman filter is constructed by combining the UKF method with the SSUT method, and it is used in the automatic astronomical navigation of spacecraft. The OSS navigation simulation experiment is carried out by using MATLAB/RTW software. The composition and implementation environment of the system, as well as some critical technologies involved, are also introduced in detail. The digital simulation of the SSUKF algorithm proves that the algorithm has the same filtering effect as the ordinary UKF algorithm. The calculation speed of the algorithm is greatly improved. A novel Kalman filter is obtained by combining the advantages of the two methods. The algorithm has good estimation accuracy, convergence speed and robustness, and higher operational efficiency. This method is suitable for autonomous positioning of spacecraft.

## REFERENCES

- [1] Cui, D. (2022). Application of CREAM algorithm in navigation. *Academic Journal of Science and Technology*, 1(2), 52-55.
- [2] Shizhuang, W. A. N. G., Xingqun, Z. H. A. N., Yawei, Z. H. A. I., Cheng, C. H. I., & Jiawen, S. H. E. N. (2021). Highly

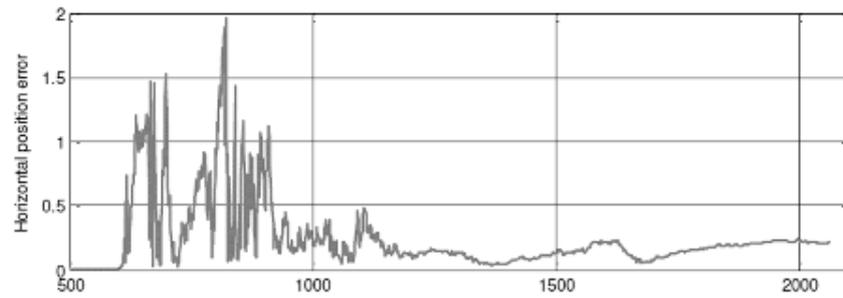


Fig. 5.2: Error estimation of positioning using SSUKF method.

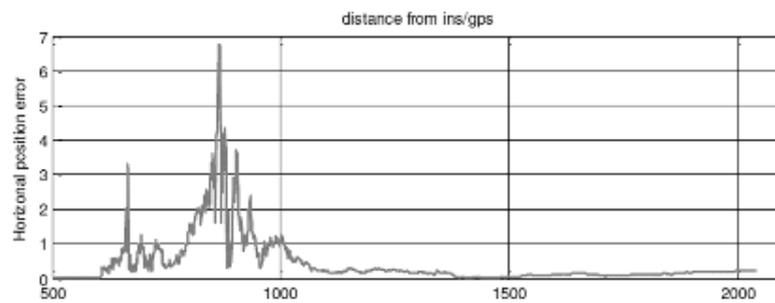


Fig. 5.3: Error analysis of speed estimation using SSUKF method.

reliable relative navigation for multi-UAV formation flight in urban environments. *Chinese Journal of Aeronautics*, 34(7), 257-270.

- [3] Faghihinia, A., Atashgah, M. A., & Dehghan, S. M. (2021). Analytical expression for uncertainty propagation of aerial cooperative navigation. *Aviation*, 25(1), 10-21.
- [4] Valasek, J., Harris, J., Pruchnicki, S., McCrink, M., Gregory, J., & Sizoo, D. G. (2020). Derived angle of attack and sideslip angle characterization for general aviation. *Journal of Guidance, Control, and Dynamics*, 43(6), 1039-1055.
- [5] Chandrasekaran, K., Theningaledathil, V., & Hebbar, A. (2021). Ground based variable stability flight simulator. *Aviation*, 25(1), 22-34.
- [6] Krasuski, K., Ćwiklak, J., Bakula, M., & Mroziak, M. (2022). Analysis of the Determination of the Accuracy Parameter for Dual Receivers Based on EGNOS Solution in Aerial Navigation. *acta mechanica et automatica*, 16(4), 365-372.
- [7] Erkec, T. Y., & HACIZADE, C. (2020). Relative navigation in UAV applications. *International Journal of Aviation Science and Technology*, 1(02), 52-65.
- [8] Chernomorsky, A. I., Lelkov, K. S., & Kuris, E. D. (2020). About One way to increase the accuracy of navigation system for ground wheeled robot used in aircraft parking. *Smart Science*, 8(4), 219-226.
- [9] Tang, Y., Zhu, F., & Chen, Y. (2021). For more reliable aviation navigation: Improving the existing assessment of airport electromagnetic environment. *IEEE Instrumentation & Measurement Magazine*, 24(4), 104-112.
- [10] Zhang, L., Hsu, L. T., & Zhang, T. (2022). A novel INS/USBL integrated navigation scheme via factor graph optimization. *IEEE Transactions on Vehicular Technology*, 71(9), 9239-9249.
- [11] Luo, W., Zhang, X., & Liu, X. (2020). Simulation of multi-path suppression algorithm for civil aviation satellite navigation jamming signal. *Computer Simulation*, 37(03), 52-56.
- [12] Sun, Y., Geng, N., Gong, S., & Yang, Y. (2020). Research on improved genetic algorithm in path optimization of aviation logistics distribution center. *Journal of Intelligent & Fuzzy Systems*, 38(1), 29-37.
- [13] Won, D., Oh, J., Kang, W., Eom, S., Lee, D., Kim, D., & Han, S. (2022). Flight Scenario Trajectory Design of Fixed Wing and Rotary Wing UAV for Integrated Navigation Performance Analysis. *Journal of the Korean Society for Aviation and Aeronautics*, 30(1), 38-43.
- [14] Wang, S. Z., Zhan, X. Q., Zhai, Y. W., Chi, C., & Liu, X. Y. (2020). Ensuring high navigation integrity for urban air mobility using tightly coupled GNSS/INS system. *J. Aeronaut. Astronaut. Aviat*, 52(4), 429-442.
- [15] Erkec, T. Y., & Hajiyev, C. (2020). The methods of relative navigation of satellites formation flight. *International Journal of Sustainable Aviation*, 6(4), 260-279.
- [16] Alnuaimi, M., & Perhinschi, M. G. (2021). Investigation of alternative parameters for immunity-based UAV navigation in GNSS-denied environment. *Unmanned Systems*, 9(01), 65-72.

- [17] Jheng, S. L., Jan, S. S., Chen, Y. H., & Lo, S. (2020). 1090 MHz ADS-B-based wide area multilateration system for alternative positioning navigation and timing. *IEEE sensors journal*, 20(16), 9490-9501.
- [18] Gao, D., Hu, B., Qin, F., Chang, L., & Lyu, X. (2021). A real-time gravity compensation method for INS based on BPNN. *IEEE Sensors Journal*, 21(12), 13584-13593.
- [19] Carroll, J. D., & Canciani, A. J. (2021). Terrain-referenced navigation using a steerable-laser measurement sensor. *Navigation*, 68(1), 115-134.
- [20] Maamo, M. S., Afonin, A. A., & Sulakov, A. S. (2022). Aircraft wing vibration parameters measurement system using MEMS IMUs and closed-loop optimal correction. *Aerospace Systems*, 5(3), 473-480.
- [21] Zhang, L., Shaoping, W. A. N. G., Selezneva, M. S., & Neusypin, K. A. (2022). A new adaptive Kalman filter for navigation systems of carrier-based aircraft. *Chinese Journal of aeronautics*, 35(1), 416-425.
- [22] Gabela, J., Kealy, A., Hedley, M., & Moran, B. (2021). Case study of Bayesian RAIM algorithm integrated with Spatial Feature Constraint and Fault Detection and Exclusion algorithms for multi-sensor positioning. *Navigation*, 68(2), 333-351.
- [23] Neogi, N. A. (2020). Progress in autonomy: space robotics, satellite systems, air traffic control. *Aerospace America*, 58(11), 44-44.
- [24] Hassan, T., El-Mowafy, A., & Wang, K. (2021). A review of system integration and current integrity monitoring methods for positioning in intelligent transport systems. *IET Intelligent Transport Systems*, 15(1), 43-60.

*Edited by:* Zhigao Zheng

*Special issue on:* Graph Powered Big Aerospace Data Processing

*Received:* Oct 3, 2023

*Accepted:* Oct 19, 2023



## VISISENSE: A COMPREHENSIVE IOT-BASED ASSISTIVE TECHNOLOGY SYSTEM FOR ENHANCED NAVIGATION SUPPORT FOR THE VISUALLY IMPAIRED

BHASHA PYDALA\*, T. PAVAN KUMAR† AND K. KHAJA BASEER‡

**Abstract.** The field of visually impaired assistive technology looks for novel approaches to enhance independence and navigation. In this field, systems have to reliably identify and transmit environmental data in order to facilitate visually impaired users' safe and effective navigation. Developing a sophisticated framework for assistive technology that significantly enhances visually impaired navigation is the goal of this research. Make object detection and environmental awareness more efficient, dependable, and intuitive. This study introduces VISISENSE, "A Comprehensive IoT-Based Assistive Technology System for Enhanced Navigation Support for the Visually Impaired." VISISENSE is an IoT-based system with multiple components that enhances object detection. For primary environmental sensing, a handstick with implant sensors, a visual capture and transmission unit for processing visual data, and edge computing for object detection and classification are used. The system makes use of the R-CNN global computer vision model hosted on a cloud server, Mobinet computer vision models, and Logistic Regression with Iterative Learning. VISISENSE's effectiveness is demonstrated by a performance analysis of its object detection accuracy, processing speed, resource utilization, energy consumption, latency, and false positive rate. In all of these categories, VISISENSE performs better than Smart Stick and Smart Navigation. The data includes the fastest processing time of 17 ms, the most efficient resource utilization of 41%, and object detection accuracy of up to 99% at 2 Mbps load. Across all load conditions, VISISENSE has the lowest false positive rate, energy consumption, and latency. The VISISENSE assistive technology system is developed for the visually impaired. Its excellent object detection and navigation accuracy, speed, and efficiency enhance user experience and hold the potential to increase the independence and quality of life for visually impaired people. This research contributes to a significant advancement in assistive devices—smart, responsive technologies for the visually impaired.

**Key words:** VisiSense, Artificial Intelligent, R-CNN, Cloudsim, MobileNet, IoT, Binary Classifier.

**1. Introduction.** People who are blind or visually impaired frequently have mobility issues. To address these issues, there is growing interest in creating IoT-based assistive technologies. [1]. This piece introduces "VisiSense," a state-of-the-art navigation aid. The objective of VisiSense is to develop a simple navigation system for those who are blind. VisiSense provides complete support and real-time assistance to visually impaired users as they navigate foreign environments. A network of devices that gathers and evaluates environmental data includes wearable sensors, smartphones, and environmental sensors. For accurate navigation assistance, VisiSense uses edge computing, cloud computing, and AI algorithms for advanced object detection and recognition.

The accuracy of object detection, processing speed, resource utilization, energy efficiency, latency, and false positive rates are all areas where VisiSense's performance analysis reveals promising results.

Due to the system's precise and timely obstacle information, visually impaired people can now confidently navigate their surroundings.

VisiSense, a navigation aid for people with visual impairments, advances this field through IoT, edge computing, cloud computing, and AI. The design, implementation, and evaluation of this comprehensive assistive navigation system demonstrate its potential to increase the freedom, security, and quality of life of blind and visually impaired people. VisiSense can increase inclusivity and accessibility for people who are blind or visually impaired.

---

\*Research Scholar, Department of CSE, KLEF (Deemed to be University), Guntur, India -522 302 (Corresponding author, [basha.chanti@gmail.com](mailto:basha.chanti@gmail.com))

†Professor, Department of CSE, KLEF (Deemed to be University), Guntur, India -522 302 ([pavankumar\\_ist@kluniversity.in](mailto:pavankumar_ist@kluniversity.in)).

‡Professor, Department of Data Science, Mohan Babu University (Erstwhile Sree Vidyanikethan Engineering College), Tirupati, India-517102 ([baseer.kk@vidyanikethan.edu](mailto:baseer.kk@vidyanikethan.edu)).

**2. Literature Survey.** This review looks at studies that suggest novel ways to improve navigation for people with vision impairments, ultimately raising their quality of life. Each paper examines a different topic and technology, highlighting its advantages and potential for development.

Islam et al.'s [1] discussion of walking assistants includes systems based on sensors, computer vision, and smartphones. In developing assistive tools, the authors stress the value of reliability, lightweight design, and simplicity. The article thoroughly examines the benefits and drawbacks of various methodologies, emphasizing the demand for further development of current systems.

A visually impaired person can use an AI-based assistive technology for smart navigation, according to Joshi et al. [2]. They use a deep learning model in their system to identify objects and provide immediate auditory feedback. The system shows promising results, achieving high accuracy rates for object detection and recognition. Notably, the addition of a distance-measuring sensor improves its performance.

Using ultrasonic sensors for obstacle detection, Saisubramanyam et al. [3] present a cheap and lightweight smart walking stick. Within a 2-meter radius, the stick can detect obstacles, including water and dark areas. The advantages of this system are its low cost, compact design, precise detection abilities, and distinctive features. The paper, however, could benefit from more in-depth programming explanations and system comparisons.

Vaidya et al. [4] develop a real-time object detection system for visually challenged individuals using image processing and machine learning techniques. The system successfully detects objects and notifies blind users through audio output. Noteworthy advantages include independent navigation for visually impaired individuals and real-time object detection. However, no specific disadvantages are mentioned in the paper.

Ramesh et al. [5] propose a low-cost walking stick for blind people equipped with ultrasonic sensors and an SMS message system for emergency situations. The stick detects obstacles within a 4-meter range and provides voice alerts. This navigation aid stands out for its affordability, efficient design, artificial vision capabilities, portability, and the inclusion of an SMS message system. No disadvantages are mentioned.

Farooq et al. [6] design an IoT-enabled smart stick incorporating ultrasonic sensors, a water sensor, a high-definition camera, and object recognition capabilities. The stick offers obstacle detection and recognition modes, providing vibration and voice feedback to users. Additionally, GPS/GSM modules enable live location tracking and emergency assistance. The system's merits include obstacle detection and recognition, voice feedback, live location tracking, and energy efficiency. No specific demerits are mentioned.

Kuriakose et al. [7] review smart navigation solutions for blind and visually impaired individuals, encompassing various tools and technologies. The article highlights the core features lacking in existing navigation systems and offers recommendations for future design. Despite the development of numerous assistive navigation technologies, there remains a need for a seamless system that functions effectively both indoors and outdoors, enables dynamic interactions, and adapts to changes.

Bhasha et al. [15], S. Maidenbaum et al. [30] and K. Patil et al. [25] designed a smart navigation stick through sensor technology. Kim et al. [16] proposed a model for improving visually impaired people experience through deep neural networks. Rahman et al. [17] proposed one IoT based smart device through IoT and Deep Learning techniques. Bhavani et al. [18] and Ashraf et al. [19] suggested a Smart Walking Stick through machine learning approach. Kim et al. [20] and Marzec et al. [21] build a healthcare device with optimal power consumption navigation system for blind people through infrared sensors.

Salama et al. [22] and Sangpal et al. [23] Proposal of Smart Stick for blind People through Arduino and AIML. Elmannai et al. [24, 26, 28] proposed an extremely precise and consistent data synthesis Framework through sensor based 6technology for Supervisory the Visually Impaired people. G.F. Lourenço et al. [27] designed an assistive device for VIP based on user satisfaction. Rajapandian et al. [29] proposed an innovative tactic AID for blind, deaf and dumb people. [31] Foley A et al. presented a review on Technology for people, not disabilities: ensuring access and inclusion.

D. Dakopoulos et al. [32] and Topknot et al. [33] presented a survey about Wearable obstacle avoidance electronic travel aids for blind. A. Iqbal et al. [34] proposed a low cost AI vision system for visually impaired people. Chirayu Shah et al. [35] proposed a Tactile Feedback based Navigation System for VIP. L. Ran et al. [36], R. Imrie et al. [38] and Mountain G et al. [37] proposed an integrated indoor/outdoor blind navigation system and quality assistive technology services.

Scherer MJ et al. [39] done a review of Living in the State of Stuck: How Technology Impacts the Lives of

Table 2.1: The Contemporary Contributions and Proposed Framework VisiSense

Reference	Using IoT	Objects on Different Sides	Edge Computing	Cloud Computing	Using AI
Joshi, Rakesh Chandra, et al. [2]	✓			✓	
Saisubramanyam, Kaladindi et al. [3]	✓				
Vaidya, Sunit et al. [4]				✓	
Ramesh, Gopisetty et al. [5]	✓				
Farooq, Muhammad Siddique et al. [6]	✓	✓			
VisiSense	✓	✓	✓	✓	✓

People with Disabilities. Md. Milon Islam et al. [40] presented a review on Development of Walking Assistants for VIP. Sreenu Ponnada et al. [41] designed a smart system for detecting Manhole and Staircase through sensor based technology. Nur Syazreen Ahmad et al. [42] developed a Multi-sensor Obstacle Detection System via Model-based State-Feedback Control in Smart Cane Design for VIP. J. Bai et al. [43] simulated a smart system model through Virtual-Blind-Road Following-Based Wearable Navigation Device for Blind People.

In assistive technologies for the visually impaired, the review addresses deep learning models, IoT, sensor-based systems, and AI applications. Reliability, ease of use, real-time processing, and affordability are necessary for assistive devices to be effective. By going over several models, the review draws attention to the expanding usage of cutting-edge technologies like AI and IoT in navigation aids. VISISENSE, which combines IoT, edge computing, cloud computing, and AI, is supported in this environment. The need for seamless indoor-outdoor functionality and user-centered design is met by this integration, which also suggests a strong and adaptable system. VISISENSE is a comprehensive solution that has the potential to improve assistive devices for the visually impaired and close many gaps in current technologies. The review demonstrates how, if it preserves user accessibility and practicality in real-world settings, its holistic approach might offer a better user experience.

**3. The Proposed Methodology.** The VisiSense framework is a comprehensive system designed to assist visually impaired individuals in detecting objects in their surroundings. It encompasses several key components that work together to provide object detection and alert capabilities. The Handstick with Implant Sensors serves as the primary sensing unit, detecting obstacles in the environment and relaying this information to the system. The Visual Capturing and Transmission Unit, consisting of a mobile camera and a processing unit, captures visual data from the surroundings and performs initial processing tasks. The Edge Computing component plays a crucial role in object classification and detection. It includes a binary classifier, specifically the Logistic Regression with Iterative Learning, which classifies visuals as "known" or "new". Known visuals are forwarded to the Mobinet computer vision model, which leverages its learned knowledge to detect objects and provide object details.

The Cloud Server component houses a global computer vision model, specifically the R-CNN (Region-based Convolutional Neural Network) [8], trained on images from multiple similar sources of the VisiSense prototype. New visuals captured by the mobile camera are transferred to the cloud server for processing. The global computer vision model applies its knowledge and detects objects on these visuals, extracting object details.

The functional flow within the framework starts with the implant sensors detecting obstacles and alerting the mobile camera. The camera captures visuals, which are processed by the unit and classified by the binary classifier. If classified as known, the visuals are sent to the computer vision model for object detection, and the processing unit alerts the user through a hearing aid about the detected objects. If classified as new, the binary classifier alerts the processing unit, which transfers the new visuals to the cloud server. The global computer vision model in the cloud server detects objects on the new visuals and sends the object details back to the processing unit. The processing unit, along with the computer vision model in the Edge Computing component, alerts the user through the hearing aid, providing information about the detected objects. VisiSense prototype is a system designed to assist visually impaired individuals in detecting objects in their surroundings (figure 3.1).

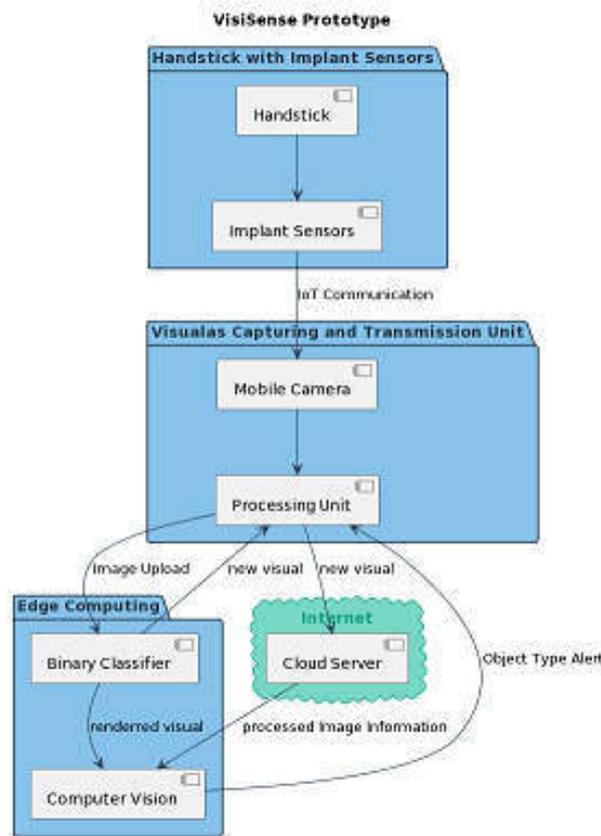


Fig. 3.1: The Flow Diagram of the Visisense Prototype Components

**3.1. Functional Flow between Components.** The implant sensors on the handstick detect obstacles in the environment and alert the mobile camera. The mobile camera captures visuals of the surroundings and transfers them to the processing unit. The processing unit forwards the visuals to the binary classifier for classification. If the visuals are classified as known, the binary classifier sends them to the computer vision model. The computer vision model detects objects on the known visuals and sends the object details back to the processing unit. The processing unit alerts the user through a hearing aid, providing information about the detected objects.

If the visuals are classified as new, the binary classifier alerts the processing unit about the new visuals and learns from them for future classification tasks. The processing unit transfers the new visuals to the cloud server.

The global computer vision model in the cloud server detects objects on the new visuals, extracting object details. The processing unit and Computer Vision Model of the Edge computing receives the object details from the cloud server and alerts the user through the hearing aid.

**3.2. The Binary classification Strategy.** Logistic regression [9] performs well and is simple for binary classification in the VisiSense prototype. Visual data are categorized as "already known / processed" (positive) or "new" using logistic regression. Logistic regression is especially suitable for the VisiSense system because it is well known for its effectiveness in handling binary classification problems. A critical function of the VisiSense framework is determining whether visual data has been processed or is new. It is well-suited for this task due to its probabilistic interpretations and adept handling of linearly separable data.

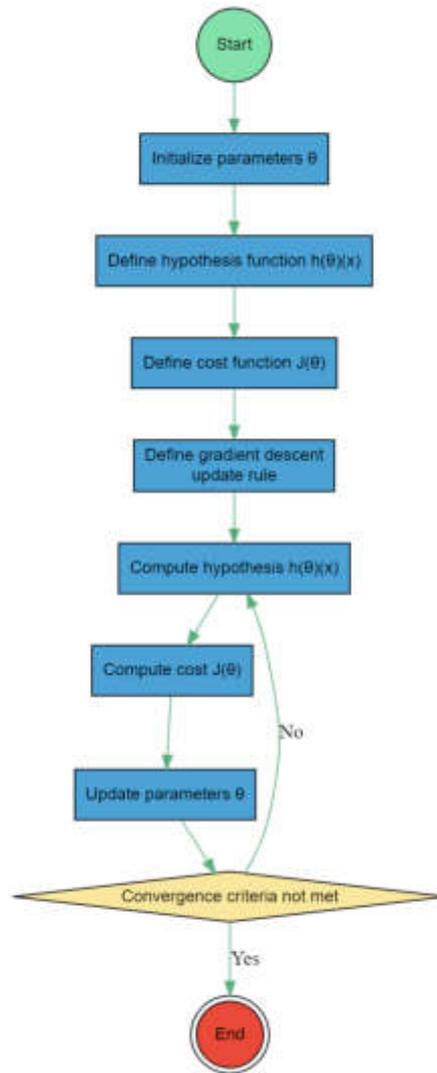


Fig. 3.2: Logistic regression algorithm

*Data Preparation.* Gather a labeled dataset of visual data, where each instance is associated with a binary label indicating whether it is "already known/processed" (positive) or "new" (negative).

Figure 3.2 describes for logistic regression algorithm with continuous learning: Represent each visual instance as a feature vector. Let  $X$  be the matrix of feature vectors, where each row corresponds to a visual instance, and let  $y$  be the vector of corresponding labels.

*Model Initialization:* Initialize the parameter vector  $\theta$  to small random values. Let  $\theta$  be a column vector with dimensions  $(n + 1) \times 1$ , where  $n$  is the number of features. The additional 1 is for the bias term.

*Hypothesis Function:* Define the hypothesis function  $h\theta(x)$  that models the probability of a visual instance belonging to the positive class as:  $h\theta(x) = \text{sigmod}(\theta^T x)$  where  $\text{sigmod}(z)$  is the sigmoid function given by Eq 3.1

$$\text{sigmod}(z) = 1 / ((1 + \exp(-z))) \quad (3.1)$$

*Cost Function.* Define the cost function  $J(\theta)$  that measures the discrepancy between the predicted probabilities and the actual labels as: Eq 3.2

$$J(\theta) = (1/m) * \text{sum}(-y * \log(h\theta(x)) - (1 - y) * \log(1 - h\theta(x))) \quad (3.2)$$

where  $m$  is the number of training examples.

*Gradient Descent.* Update the parameters iteratively using gradient descent to minimize the cost function is Eq 3.3.

$$J(\theta) : \theta = \theta - * (1/m) * X^T * (h\theta(x) - y) \quad (3.3)$$

*Continuous Learning.* During the system's operation, continuously monitor the predictions and feedback received from false positives, false negatives, and true negatives.

For each misclassification, update the parameter vector  $\theta$  by further adjusting the parameters using the gradient descent update rule.

The logistic regression algorithm optimizes the parameters  $\theta$  to maximize the likelihood of the observed labels given the input visual features. By iteratively updating the parameters based on the gradient descent algorithm and continuously learning from misclassifications, the model improves its ability to classify "already known/processed" and "new" visuals more accurately.

*Mathematical Model.* The logistic regression algorithm can be formulated as follows.

By iteratively updating the parameters using gradient descent and continuously learning from misclassifications, the logistic regression model in the VisiSense prototype adapts to new visual data, improving its classification performance over time.

1. Hypothesis Function:  $h\theta(x) = \text{sigmoid}(\theta^T x)$
2. Cost Function:  $J(\theta) = (1/m) * \text{sum}(-y * \log(h\theta(x)) - (1 - y) * \log(1 - h\theta(x)))$
3. Gradient Descent Update Rule:  $\theta := \theta - * (1/m) * X^T * (h\theta(x) - y)$

In the above equations,  $X$  represents the matrix of feature vectors,  $y$  represents the vector of labels,  $\theta$  represents the parameter vector, and  $*$  represents the learning rate. The sigmoid function is given by Eq 3.4.

$$\text{sigmoid}(z) = (1/((1 + \exp(-z)))) \quad (3.4)$$

By iteratively updating the parameters using gradient descent and continuously learning from misclassifications, the logistic regression model in the VisiSense prototype adapts to new visual data, improving its classification performance over time.

**3.3. Computer Vision Model as Edge Computing Service .** The constrained resource environment of the VisiSense prototype is ideal for MobileNet's [10] compact convolutional neural network (CNN) architecture. Without sacrificing precision, depth-wise separable convolutions reduce the number of parameters and computational complexity. MobileNet allows for customization based on the size of the model and the required computational power. Real-time processing and user alerts are feasible due to the system's lightweight design. Transfer learning improves performance with limited labeled data in MobileNet. Because of its efficacy, adaptability, and real-time processing, MobileNet is ideal for the VisiSense prototype's efficient and accurate object recognition, which improves the perception and safety of the blind and visually impaired (see Figure 3.3).

The steps involved in object detection model follows:

*Step 1. Model Initialization and Configuration.* Load the MobileNet architecture with a specific configuration suitable for the VisiSense prototype, considering the balance between model size, accuracy, and computational efficiency.

*Step 2. Input Preprocessing.* Preprocess the visual data captured by the wearable mobile camera to meet the input requirements of MobileNet. This typically involves resizing the images to the input dimensions expected by MobileNet and normalizing the pixel values.

*Step 3. Convolutional Layers.* Pass the preprocessed visual data through the convolutional layers of MobileNet, which consist of depth-wise separable convolutions. For each layer, compute the weighted sum of inputs and apply a non-linear activation function, such as ReLU: a. Depth-wise Convolution:

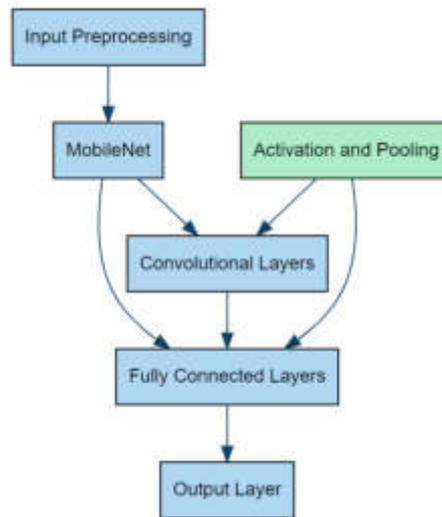


Fig. 3.3: The MobileNet for VisiSense

1. Perform depth-wise convolution for each channel  $i$ :  $z_i = W_i * x$ : Apply the activation function (e.g.,  $ReLU$ ):  $a_i = ReLU(z_i)$ .
- 2.
3. Point-wise Convolution:
  - (a) Perform point-wise convolution across channels:  $z = W * a$
  - (b) Apply the activation function:  $a = ReLU(z)$

*Step 4: Activation and Pooling.* Apply non-linear activation functions, such as  $ReLU$ , after each convolutional layer to introduce non-linearity into the model and capture complex features. Perform down sampling using pooling layers, such as max pooling or average pooling, to reduce the spatial dimensions of the feature maps while preserving important information.

*Step 5: Fully Connected Layers.* Flatten the output feature maps from the convolutional layers into a 1D vector. Connect the flattened vector to one or more fully connected layers, which process the extracted features and capture higher-level representations. Apply an activation function, such as  $ReLU$ , to the outputs of the fully connected layers.

*Step 6: Output Layer.* Connect the output of the fully connected layers to the final output layer. Apply an appropriate activation function based on the task, such as softmax for multi-class classification or sigmoid for binary classification. Obtain the predicted probabilities or class labels from the output layer.

**3.4. Mathematical Model.** Let's denote the input visual data as  $x$ , the output probabilities as  $y_{pred}$ , and the parameters of MobileNet as  $W_i$ ,  $W$  and  $W_{fc}$ .

*Convolutional Layers.* For each layer, compute the weighted sum of inputs and apply the activation function:

1. Depth-wise Convolution:
  - (a) Perform depth-wise convolution for each channel  $i$ :  $z_i = W_i * x$ .
  - (b) Apply the activation function (e.g.,  $ReLU$ ):  $a_i = ReLU(z_i)$ .
2. Point-wise Convolution:
  - (a) Perform point-wise convolution across channels:  $z = W * a$ .
  - (b) Apply the activation function:  $a = ReLU(z)$ .

*Activation and Pooling.* Apply non-linear activation functions, such as  $ReLU$ , after each convolutional layer. Perform down sampling using pooling layers, such as max pooling or average pooling.

*Fully Connected Layers.*

1. Flatten the output feature maps into a 1D vector:  $a_{flat} = Flatten(a)$ .

2. Compute the weighted sum of the flattened vector:  $z_{fc} = W_{fc} * a_{flat}$ .
3. Apply the activation function:  $a_{fc} = \text{ReLU}(z_{fc})$ .

*Output Layer.*

1. Compute the weighted sum of the fully connected layer output:  $z_{out} = W_{out} * a_{fc}$ .
2. Apply the appropriate activation function based on the task (e.g., softmax for multi-class classification or sigmoid for binary classification):  $y_{pred} = \text{activation}(z_{out})$ .

The parameters  $W_i$ ,  $W$ ,  $W_{fc}$ , and  $W_{out}$  represent the weights of the respective layers, and ReLU is the activation function used throughout the model.

By following these algorithmic steps and applying the mathematical model, MobileNet within the VisiSense prototype efficiently processes visual data captured by the wearable mobile camera. It produces predicted probabilities or class labels, indicating the presence and type of objects within the captured visual data. The lightweight and efficient nature of MobileNet allows for real-time object identification on edge computing devices, enhancing the accuracy and responsiveness of the VisiSense system for the benefit of visually impaired individuals.

**3.5. Global Computer Vision Model as Cloud Service.** The R-CNN (Region-based Convolutional Neural Network) [8] algorithm holds significant significance as the Global Computer Vision Model in the VisiSense prototype. Figure 3.4 R-CNN provides accurate and robust object detection by employing a two-stage approach, with a region proposal network (RPN) [11] generating potential object proposals and a subsequent classification network performing object classification. This two-stage process ensures precise localization and classification of objects in visual data. The RPN efficiently generates region proposals using anchor boxes and convolutional features, enabling the model to handle objects of various sizes and aspect ratios. Additionally, R-CNN offers flexibility, adaptability, and real-time processing capabilities. Its popularity in the computer vision community provides access to pre-trained models, resources for fine-tuning on specific datasets, and an active support community, making it a valuable choice for enhancing object detection in the VisiSense prototype and empowering visually impaired individuals with improved situational awareness. R-CNN (Region-based Convolutional Neural Network) is a two-stage algorithm for object detection. It consists of a region proposal network (RPN) and a subsequent classification network. Here's a high-level description of the R-CNN algorithm and its mathematical model.

**3.5.1. Region Proposal Network (RPN).** The RPN generates region proposals by sliding a small network (typically a convolutional neural network) over the input image. It predicts potential object bounding boxes, known as region of interests (RoIs), and their associated objectness scores. The RPN operates on a set of anchor boxes of different scales and aspect ratios, which act as reference boxes for proposing potential object locations. The RPN predicts the objectness score (probability of an anchor containing an object) and the coordinates ( $x$ ,  $y$ , width, height) of the bounding boxes relative to the anchors.

**Inputs:** Image ( $I$ )

Anchor boxes:  $A = \{a_1, a_2, \dots, a_n\}$  (n anchor boxes with different scales and aspect ratios)

Output: Region proposals ( $R$ ), Objectness scores ( $O$ )

RPN predicts the objectness score for each anchor box:

$O = \{o_1, o_2, \dots, o_n\}$  (where  $o_i$  is the objectness score of the  $i$ -th anchor)

RPN predicts the coordinates for each anchor box:  $R = \{r_1, r_2, \dots, r_n\}$  (where  $r_i = (dx_i, dy_i, dw_i, dh_i)$  represents the bounding box coordinates relative to the  $i$ -th anchor)

**3.5.2. Classification Network.** The classification network takes the proposed RoIs from the RPN and extracts fixed-sized feature vectors using a shared convolutional backbone network.

These feature vectors are then fed into a classifier (typically a fully connected network) to classify the object contained within each proposed RoI. The classifier predicts the object class probabilities and, optionally, refines the bounding box coordinates of the proposed RoIs.

**Inputs:** Proposed RoIs ( $R'$ )

**Output:** Class probabilities ( $P$ ), Refined bounding box coordinates ( $B$ )

The classification network takes the proposed RoIs  $R'$  and extracts feature vectors using a shared convolutional backbone network.

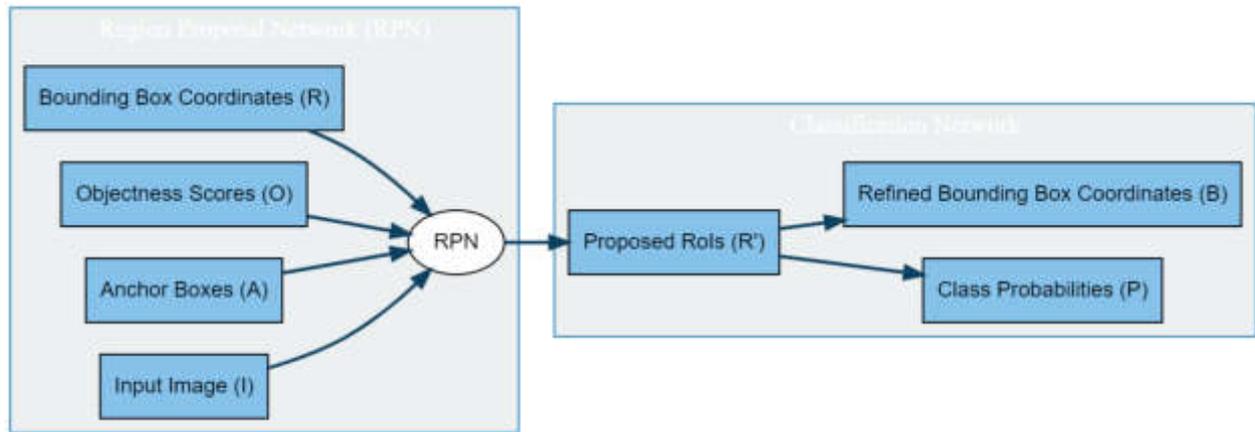


Fig. 3.4: R-CNN for global computer vision model

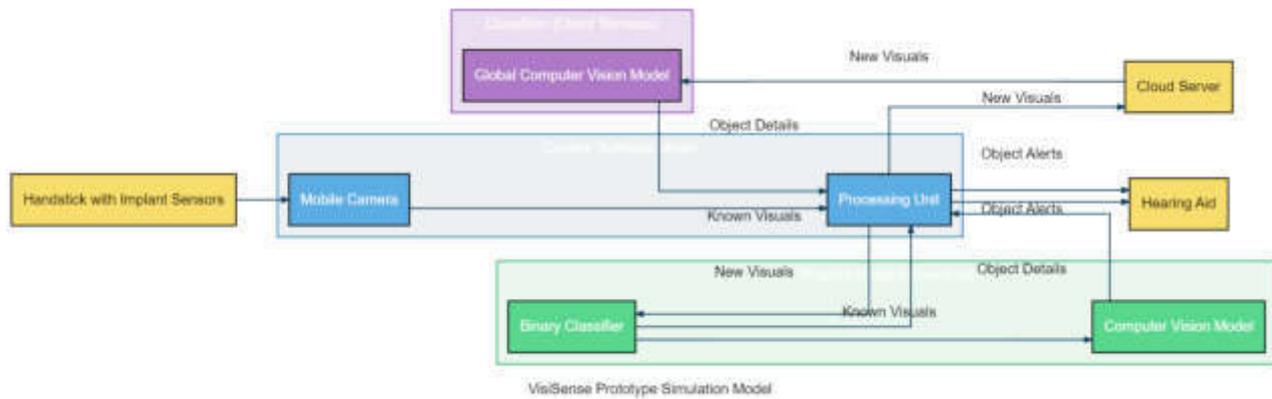


Fig. 4.1: VisiSense prototype simulation model

The feature vectors are fed into a classifier to predict the class **probabilities**:  $P = \{p_1, p_2, \dots, p_m\}$  (where  $p_i$  is the probability of the  $i$ -th RoI belonging to a specific object class).

The classifier may also refine the bounding box coordinates of the proposed RoIs:  $B = \{b_1, b_2, \dots, b_m\}$  (where  $b_i = (dx'_i, dy'_i, dw'_i, dh'_i)$  represents the refined bounding box coordinates of the  $i$ -th RoI).

**4. Experimental Study.** The Simulation Model depicted in figure 4.1 is defined to test the performance of VisiSense. The simulation model that integrates Custom Software Model, iFogSim [12] for Edge Computing, and CloudSim [13] for Cloud Services. The Custom Software Model comprises the "Mobile Camera" and "Processing Unit" components. The "Mobile Camera" captures visuals from the environment, which are then processed by the "Processing Unit." The processed visuals are further passed to the "Binary Classifier" for classification, and the resulting visual classification is sent to the "Hearing Aid" for generating object alerts.

In Figure 4.1, the iFogSim module represents the Edge Computing component of the VisiSense Prototype Simulation Model, playing a critical role by enabling low-latency, real-time processing crucial for assisting visually impaired users. Edge Computing's proximity to data sources allows for rapid processing and decision-making, enhancing user safety through immediate object recognition and alerts. It reduces the system's reliance on cloud connectivity, ensuring consistent performance even with variable network quality.

This local processing capability not only optimizes bandwidth and energy efficiency, crucial for handheld assistive devices, but also bolsters privacy by minimizing the transmission of sensitive data. Furthermore, the Edge Computing layer, through its Binary Classifier, provides a scalable and adaptive mechanism that learns from new visual data, thereby continuously refining the system's accuracy and responsiveness. Overall, Edge Computing significantly contributes to the VisiSense system's robustness, making it an indispensable element in the proposed research work for real-time navigational support to visually impaired individuals.

The "Binary Classifier," a component of the iFogSim package, examines the images received from the "Processing Unit" to determine whether they are old or new images. The "Computer Vision Model" then processes known images for object detection, sending the resulting object details back to the "Processing Unit."

When there are new visuals, the "Binary Classifier" alerts the "Processing Unit" to their presence. The "Processing Unit" then sends these images to the "Global Computer Vision Model" inside the CloudSim package, which carries out object detection. The "Processing Unit" and the "Computer Vision Model" are both given access to the object details discovered from the fresh visuals.

The "Processing Unit" creates object alerts by combining the object details obtained from the "Global Computer Vision Model" and the "Computer Vision Model." Users can learn more about the discovered objects thanks to the "Hearing Aid [14]," which receives these alerts.

**4.1. Performance Analysis.** By examining important performance indicators like object detection accuracy, processing speed, resource utilization, energy consumption, latency, and false positive rate, this section evaluates VisiSense. VisiSense proved its superiority and efficacy. Accurate object detection, real-time data processing, resource optimization, energy efficiency, and low false positive rates are just a few of its strong points. The section also contrasts VisiSense with other current models to show its benefits and potential influence on the assistive technology market. The performance analysis aids in decision-making and ongoing development while showcasing VisiSense's capabilities and directing improvements. In the end, VisiSense shows promise in enhancing accessibility and quality of life for people with visual impairments.

**4.1.1. Object Detection Accuracy.** This metric evaluates the precision of VisiSense's object detection. It evaluates the precision of object identification in the recorded images. It's crucial to evaluate both overall accuracy and object type accuracy in order to gauge the system's object detection capabilities.

The Figure 4.2, Table 4.1 shows the results for "Object Detection Accuracy". The values in Mbps in the table compare the object detection accuracy of VisiSense, Smart Stick [6], and Smart Navigation [7] under various load conditions.

VisiSense achieves the highest object detection accuracy of 99% at a load of 2 Mbps. The Smart Stick comes in second place at 91% and Smart Navigation comes in at 84%. VisiSense maintains a high accuracy of 97% as the load increases to 4 Mbps, while the Smart Stick and Smart Navigation achieve 90% and 83% accuracy, respectively. VisiSense consistently outperforms the other solutions in terms of object detection accuracy as the load increases, and this trend continues.

VisiSense achieves 94% accuracy at 10 Mbps, compared to 88% and 82% accuracy for the Smart Stick and Smart Navigation, respectively. As the load grows, accuracy gradually declines, but VisiSense consistently keeps its accuracy higher than the other two solutions.

**4.1.2. Processing Time.** In this metric, VisiSense analyzes visual data and identifies objects. It takes time to capture visuals and to perform preprocessing, classification, and object detection. People with visual impairments require quicker responses and real-time performance Figure 4.3.

Table 4.2 shows the processing time in milliseconds (ms) for VisiSense, Smart Stick, and Smart Navigation under various load conditions, denoted by the values in Mbps.

VisiSense displays the quickest processing time of 17 ms at a load of 2 Mbps, followed by the Smart Stick at 26 ms and Smart Navigation at 32 ms. The processing time for VisiSense slightly increases to 19 ms at 4 Mbps load, while the processing times for the Smart Stick and Smart Navigation are 28 ms and 34 ms, respectively. As the load continues to rise, this pattern persists.

VisiSense processes data at a rate of 24 ms at 10 Mbps, compared to 35 ms and 40 ms for the Smart Stick and Smart Navigation, respectively. As the load increases, processing time gradually increases, but VisiSense consistently shows faster processing times than the other two models. Overall, these findings suggest that

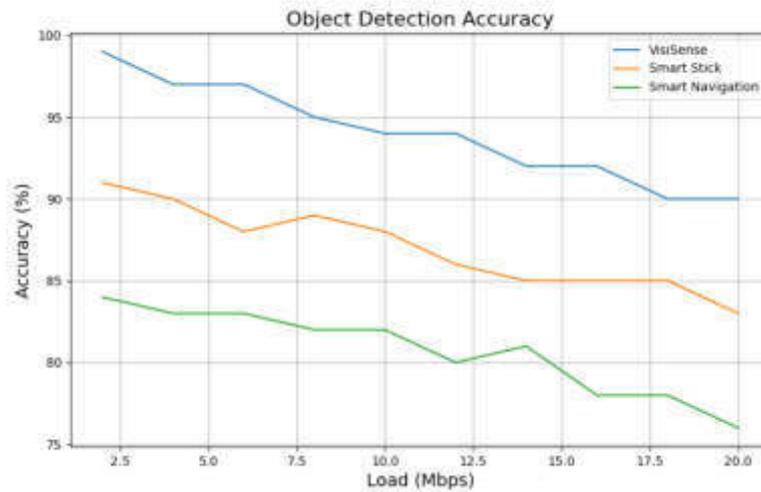


Fig. 4.2: The Object detection accuracy for VisiSense, Smart Stick, and Smart Navigation

Table 4.1: Comparison of Performance Metrics at Different Loads

Load (Mbps)	VisiSense	Smart Stick	Smart Navigation
2	99%	91%	84%
4	97%	90%	83%
6	97%	88%	83%
8	95%	89%	82%
10	94%	88%	82%
12	94%	86%	80%
14	92%	85%	81%
16	92%	85%	78%
18	90%	85%	78%
20	90%	83%	76%

VisiSense processes information more quickly than the Smart Stick and Smart Navigation under various load conditions. This suggests that VisiSense can process information more quickly and effectively, leading to faster response times. Although slightly slower than VisiSense, the Smart Stick and Smart Navigation also display acceptable processing times.

These results highlight how quickly and effectively VisiSense processes information, making it a promising assistive technology for people who are blind or visually impaired. The system is more usable overall because of the quicker object detection and recognition and more seamless user experience that results from the shorter processing time.

**4.1.3. Resource Utilization.** This metric evaluates the efficiency with which VisiSense makes use of the CPU, memory, and network bandwidth. It assesses the effectiveness of visual processing and resource allocation. Cost and efficiency are both increased by using fewer resources.

The Figure 4.4, Table 4.3 representation of the resource utilization for VisiSense, Smart Stick, and Smart Navigation in percentages (%) for various load conditions.

VisiSense shows a resource utilization of 41% at a load of 2 Mbps, compared to resource utilizations of 50% and 42% for the Smart Stick and Smart Navigation, respectively. VisiSense's resource utilization rises to 47% at 4 Mbps, while the resource utilization for the Smart Stick and Smart Navigation is 57% and 54%,

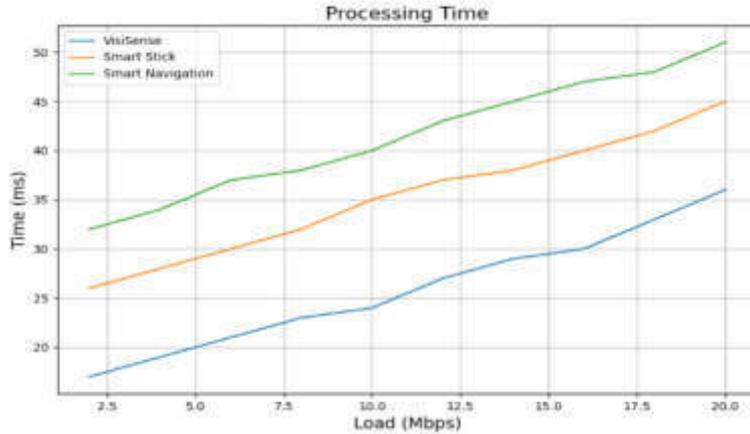


Fig. 4.3: The processing time in milliseconds (ms) for VisiSense, Smart Stick, and Smart Navigation

Table 4.2: Response Time Analysis Across Different Systems

Load (Mbps)	VisiSense	Smart Stick	Smart Navigation
2	17	26	32
4	19	28	34
6	21	30	37
8	23	32	38
10	24	35	40
12	27	37	43
14	29	38	45
16	30	40	47
18	33	42	48
20	36	45	51

respectively. As the load continues to rise, this pattern persists.

VisiSense uses 58% of the resources at 10 Mbps, while the Smart Stick and Smart Navigation use 68% and 66% of the resources, respectively. As the load increases, resource utilization gradually rises, with VisiSense consistently showing lower resource utilization than the other two models.

These findings collectively show that VisiSense uses resources more sparingly than the Smart Stick and Smart Navigation under various load conditions. This suggests that VisiSense makes better use of system resources through optimization, which enhances performance and stability. While slightly higher than VisiSense, the Smart Stick and Smart Navigation also show acceptable resource usage.

These results demonstrate VisiSense’s effectiveness and resourcefulness, making it a viable and useful assistive technology for people who are blind or visually impaired. Lower resource utilization guarantees improved system performance and scalability, enabling more seamless operation and better user experience.

**4.1.4. Energy Consumption.** This metric assesses the energy required for processing visual data by VisiSense. It takes into account the processing power, edge computing hardware, cloud servers, and mobile camera’s energy consumption.

Energy optimization reduces system energy requirements and increases the battery life of mobile devices Figure 4.5, Table 4.4. The provided table shows VisiSense, Smart Stick, and Smart Navigation energy consumption in millijoules for various load conditions, represented by the values in Mbps.

VisiSense uses 127 millijoules of energy at a load of 2 Mbps, compared to 182 millijoules and 151 millijoules

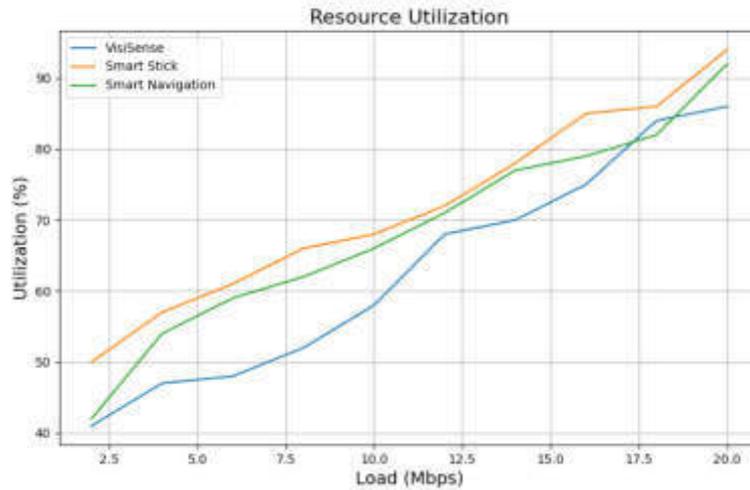


Fig. 4.4: Resource utilization for VisiSense, Smart Stick, and Smart Navigation in percentages (%) for various load conditions

Table 4.3: Comparison of Load Distribution Efficiency

Load (Mbps)	VisiSense	Smart Stick	Smart Navigation
2	41%	50%	42%
4	47%	57%	54%
6	48%	61%	59%
8	52%	66%	62%
10	58%	68%	66%
12	68%	72%	71%
14	70%	78%	77%
16	75%	85%	79%
18	84%	86%	82%
20	86%	94%	92%

used by the Smart Stick and Smart Navigation, respectively. VisiSense's energy consumption increases to 152 millijoules at 4 Mbps load, while the Smart Stick and Smart Navigation use 206 milli-joules and 182 milli-joules of energy, respectively. As the load rises even higher, this pattern persists.

VisiSense uses 228 millijoules of energy at 10 Mbps, compared to 280 millijoules and 257 millijoules for the Smart Stick and Smart Navigation, respectively. As the load increases, the energy consumption rises gradually, with VisiSense consistently showing lower energy consumption than the other two models.

These findings show that VisiSense uses considerably less energy than the Smart Stick and Smart Navigation under various load conditions. This implies that VisiSense is more energy-efficient and optimized, resulting in lower power usage and longer battery life. Both the Smart Stick and Smart Navigation exhibit reasonable energy usage, albeit a little bit more so than VisiSense.

These findings highlight VisiSense's sustainability and energy efficiency, making it an appealing assistive technology for people who are blind or visually impaired. Because of the lower energy consumption, devices can run for longer periods of time and require fewer battery changes or recharges overall. In the end, VisiSense provides a dependable and cost-effective solution to increase the independence and mobility of people with visual impairments.

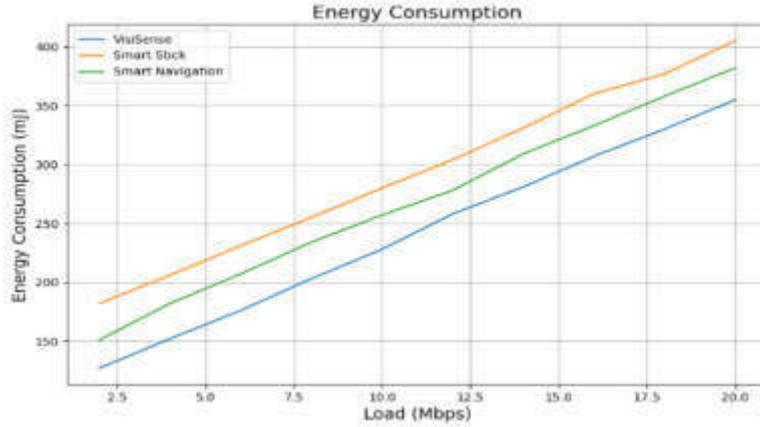


Fig. 4.5: Energy Consumption for VisiSense, Smart Stick, and Smart Navigation

Table 4.4: Performance Metrics Across Different Load Levels

Load (Mbps)	VisiSense	Smart Stick	Smart Navigation
2	127	182	151
4	152	206	182
6	176	231	207
8	203	255	234
10	228	280	257
12	258	304	278
14	281	331	309
16	307	360	333
18	330	377	358
20	355	405	382

The VisiSense system is compared at a 10Mbps data rate. In order to balance quick data transfer with the bandwidth available in locations where such a system could be used, this data rate was selected. If you want to know how well the system works in scenarios that users might encounter frequently, the 10Mbps rate is a good place to start. Figure 4.5 shows the processing efficiency of VisiSense at 10Mbps, showing how fast it can process large amounts of data in real time—a critical feature for applications requiring users of assistive technology. The system’s response time is displayed by this rate, which is helpful in circumstances that change quickly. Figure 4.6 displays system optimization and resource utilization at 10Mbps. Compared to similar systems, VisiSense uses CPU, memory, and network bandwidth more efficiently. Maintaining the system’s smooth and efficient operation depends heavily on resource allocation efficiency. Figure 4.7 illustrates the system’s lower energy consumption compared to other models for VisiSense at 10Mbps. Energy efficiency improves user experience for devices that are used for extended periods of time without being charged.

**4.1.5. Latency.** This metric assesses the communication overhead of VisiSense’s handstick, mobile camera, processing unit, edge computing devices, and cloud servers.

The provided Figure 4.6, Table 4.5 displays the VisiSense, Smart Stick, and Smart Navigation latency in milliseconds for various load conditions, represented by the values in Mbps. VisiSense exhibits a latency of 13 ms at a load of 2 Mbps, compared to latencies of 33 ms and 24 ms for the Smart Stick and Smart Navigation, respectively. The latency for VisiSense increases to 16 ms at 4 Mbps of load, while the latencies for the Smart Stick and Smart Navigation are 42 ms and 29 ms, respectively. As the load continues to rise, this pattern

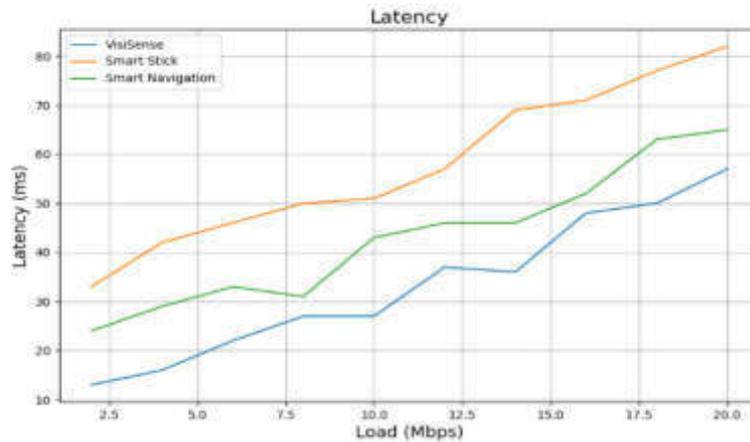


Fig. 4.6: The VisiSense, Smart Stick, and Smart Navigation latency in milliseconds for various load conditions

Table 4.5: Comparative Analysis of Load Distribution

Load (Mbps)	VisiSense	Smart Stick	Smart Navigation
2	13	33	24
4	16	42	29
6	22	46	33
8	27	50	31
10	27	51	43
12	37	57	46
14	36	69	46
16	48	71	52
18	50	77	63
20	57	82	65

persists.

The Smart Stick and Smart Navigation have latencies of 51 milliseconds and 43 milliseconds, respectively, while VisiSense has a latency of 27 milliseconds at 10 Mbps. As the load rises, the latency gradually increases, but VisiSense consistently maintains a lower latency than the other two models.

In conclusion, these results show that VisiSense has lower latency than the Smart Stick and Smart Navigation under various load conditions. This suggests that VisiSense enables real-time interaction and navigation for people who are blind or visually impaired by offering quicker response times and more effective data processing. Even though they have slightly longer latencies than VisiSense, the Smart Stick and Smart Navigation also exhibit acceptable latencies. These findings demonstrate VisiSense's effectiveness as an assistive technology for people who are blind or visually impaired. Lower latency guarantees accurate and timely information delivery, improving user experience and facilitating fluid navigation. In the end, VisiSense provides a dependable and adaptable solution to increase the independence and mobility of people with visual impairments.

**4.1.6. False Positive Rate.** These metrics gauge the frequency of false positives and negatives in VisiSense's object detection. When an object is wrongly classified or overlooked, false positives and negatives are produced. Improved system dependability and trustworthiness result from fewer false positives and negatives. The Figure 4.7, Table 4.6 gives the percentage false positive rates for VisiSense, Smart Stick, and Smart Navigation under various load conditions, represented by the values in Mbps.

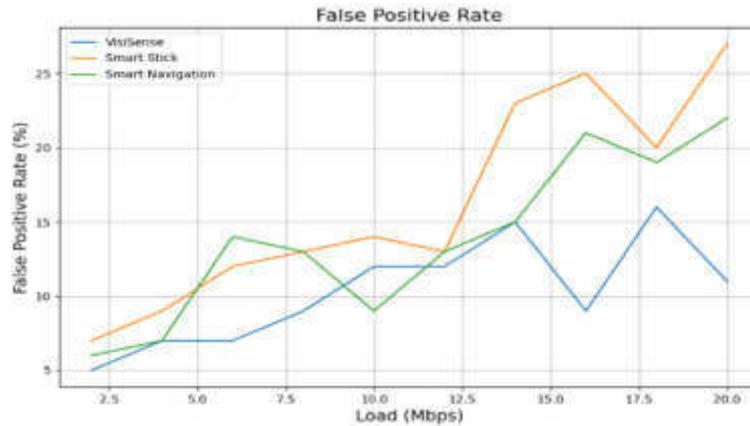


Fig. 4.7: The False positive rates for VisiSense, Smart Stick, and Smart Navigation under various load conditions

Table 4.6: Performance Metrics at Various Network Loads

Load (Mbps)	VisiSense	Smart Stick	Smart Navigation
2	5	7	6
4	7	9	7
6	7	12	14
8	9	13	13
10	12	14	9
12	12	13	13
14	15	23	15
16	9	25	21
18	16	20	19
20	11	27	22

The false positive rates of VisiSense, the Smart Stick, and Smart Navigation were evaluated across different load conditions. At a load of 2 Mbps, VisiSense exhibited a 5% false positive rate, while the Smart Stick and Smart Navigation had rates of 7% and 6%, respectively. As the load increased to 4 Mbps, VisiSense's false positive rate rose to 7%, with the Smart Stick and Smart Navigation showing rates of 9% and 7%. As the load increased further, this trend persisted at varying rates.

VisiSense's false positive rate at 10 Mbps was 12%, which was lower than that of the Smart Stick (14%) and the Smart Navigation (9%). False positive rates also fluctuated between models and loads.

The comparison as a whole demonstrates that VisiSense maintains a low false positive rate across a wide range of workloads. This shows that it detects objects more precisely and with fewer false positives than competing models. The Smart Stick and Smart Navigation performed adequately in object detection, despite having slightly higher false positive rates. Maintaining a low false positive rate is a key component of VisiSense's reliability and usability as an assistive technology, allowing it to provide visually impaired users with a trustworthy and reliable navigation solution. As a result of its ability to reduce false positives, VisiSense helps people with visual impairments move around more securely and with greater assurance.

Using the false positive rate, Figure 4.7 determines the object detection precision of the VisiSense system under different network load scenarios. False positives are situations in which the system incorrectly recognizes an object in visual data in the graph. This rate is critical to system reliability because a lower false positive rate denotes higher accuracy and reliability. VisiSense performs better than Smart Stick and Smart Navigation,

as evidenced by its false positive rate being lower under various loads.

The ratio of false positive detections to the total number of objects—true and false—is known as the false positive rate. This is a typical performance metric for object detection systems. By maintaining a lower rate of false positives, the graph in Figure 4.7 illustrates the dependability of the system and shows how effective VisiSense is at assisting visually impaired users with navigation. For users who depend on the system for safe navigation, the graph illustrates the model’s accuracy in object detection.

**5. Conclusion.** Quantitative gains over current models demonstrate that VisiSense represents a significant advancement in assistive technologies for the visually impaired. To enhance navigational aids, the system integrates cutting-edge technologies such as edge computing, cloud computing, Internet of Things, and artificial intelligence. With a 2Mbps object detection accuracy of 99%, VisiSense outperforms both Smart Stick (91%), and Smart Navigation (84%). VisiSense processes in 17 milliseconds, while the Smart Stick and Smart Navigation take 26 and 32 milliseconds, respectively. VisiSense exhibits superior efficiency by using only 41% of its available bandwidth at 2Mbps, while its competitors use 50% and 42%. The system uses 228 millijoules of energy at 10Mbps, which is significantly less than the 280 and 257 millijoules used by the Smart Stick and Smart Navigation, respectively. These figures show not only how much VisiSense has improved, but also how dependable it is—a feature that makes it a dependable tool for users. According to the study, VisiSense can significantly enhance the security, autonomy, and quality of life of people with visual impairments. With continued research and development, VisiSense has the potential to transform assistive navigation and empower people who are visually impaired.

#### REFERENCES

- [1] MD MILON ISLAM, ET AL., “Developing walking assistants for visually impaired people: A review,” *IEEE Sensors Journal*, 19.8 (2019): 2814-2828.
- [2] RAKESH CHANDRA JOSHI, ET AL., “Efficient multi-object detection and smart navigation using artificial intelligence for visually impaired people,” *Entropy*, 22.9 (2020): 941.
- [3] KALADINDI SAISUBRAMANYAM, AND A. JHANSI RANI, “IOT Based Smart Walking Stick for Blind or Old Aged People with GPS Tracking Location,” *Journal of Engineering Science*, 11.1 (2020): 501-505.
- [4] SUNIT VAIDYA, ET AL., “Real-time object detection for visually challenged people,” in *2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS)*, IEEE, 2020.
- [5] GOPSETTY RAMESH, NARENDRA B. MUSTARE, AND K. UDAY3 C. PRAMOD KUMAR, “Development of E-Stick for Blind Persons using IoT,” *I-Manager’s Journal on Embedded Systems*, 10.2 (2022).
- [6] MUHAMMAD SIDDIQUE FAROOQ, ET AL., “IoT Enabled Intelligent Stick for Visually Impaired People for Obstacle Recognition,” *Sensors*, 22.22 (2022): 8914.
- [7] BINEETH KURIAKOSE, RAJU SHRESTHA, AND FRODE EIKA SANDNES, “Tools and technologies for blind and visually impaired navigation support: a review,” *IETE Technical Review*, 39.1 (2022): 3-18.
- [8] ROSS GIRSHICK, ET AL., “Region-based convolutional networks for accurate object detection and segmentation,” *IEEE transactions on pattern analysis and machine intelligence*, 38.1 (2015): 142-158.
- [9] RAYMOND E. WRIGHT, “Logistic regression,” (1995).
- [10] HONG-YEN CHEN, AND CHUNG-YEN SU, “An enhanced hybrid MobileNet,” in *2018 9th International Conference on Awareness Science and Technology (iCAST)*, IEEE, 2018.
- [11] S. REN, ET AL., “Faster r-cnn: Towards real-time object detection with region proposal networks,” *Advances in Neural Information Processing Systems*, 28 (2015).
- [12] H. GUPTA, ET AL., “iFogSim: A toolkit for modeling and simulation of resource management techniques in the Internet of Things, Edge and Fog computing environments,” *Software: Practice and Experience*, 47.9 (2017): 1275-1296.
- [13] R. KUMAR, AND G. SAHOO, “Cloud computing simulation using CloudSim,” *arXiv preprint arXiv:1403.3253*, (2014).
- [14] B. EDWARDS, “The future of hearing aid technology,” *Trends in Amplification*, 11.1 (2007): 31-45.
- [15] P. BHASHA, DR. T. P. KUMAR, AND DR. K. K. BASEER, “A Simple and Effective Electronic Stick to Detect Obstacles for Visually Impaired People through Sensor Technology,” *Journal of Advanced Research in Dynamical & Control Systems*, 12, Issue-06, (2020).
- [16] Y. KIM, H. JEONG, J.-D. CHO, AND J. SHIN, “Construction of a Soundscape-Based Media Art Exhibition to Improve User Appreciation Experience by Using Deep Neural Networks,” *Electronics*, 2021, 10, 1170.
- [17] M. W. RAHMAN, ET AL., “The Architectural Design of Smart Blind Assistant Using IoT with Deep Learning Paradigm,” *Internet of Things*, 13, (2021), Article ID: 100344.
- [18] R. BHAVANI, ET AL., “Development of a Smart Walking Stick for Visually Impaired People,” *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, 12, (2021): 999-1005.
- [19] A. ASHRAF, ET AL., “IoT Empowered Smart Stick Assistance for Visually Impaired People,” *International Journal of Scientific & Technology Research*, 9, (2020): 356-360.

- [20] T. H. KIM, H. KOO, AND E. HAN, "Healthcare Use among Persons with Visual Impairment: Difference-in-Difference Estimations," *Disability and Health Journal*, 12(2) (2019): 302–9.
- [21] P. MARZEC AND A. KOS, "Low energy precise navigation system for the blind with infrared sensors," in *2019 MIXDES-26th International Conference Mixed Design of Integrated Circuits and Systems*, IEEE, 2019, pp. 394–397.
- [22] R. SALAMA AND A. AYOUB, "Design of Smart Stick for Visually Impaired People Using Arduino," *New Trends and Issues Proceedings on Humanities and Social Sciences*, 6 (2019): 58–71.
- [23] R. SANGPAL, T. GAWAND, S. VAYKAR AND N. MADHAVI, "JARVIS: An interpretation of AIML with integration of gTTS and Python," in *2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICT)*, Kannur, Kerala, India, 2019, pp. 486–489.
- [24] W. M. ELMANNAI AND K. M. ELLEITHY, "A Highly Accurate and Reliable Data Fusion Framework for Guiding the Visually Impaired," *IEEE Access*, vol. 6, 2018, pp. 33029–33054.
- [25] K. PATIL, Q. JAWADWALA, AND F.-L. C. SHU, "Design and construction of electronic Aid for visually impaired people," *IEEE Transactions on Human-Machine Systems*, vol. 48, no. 2, 2018, pp. 172–182.
- [26] W. ELMANNAI AND K. ELLEITHY, "Sensor-based assistive devices for visually-impaired people: current status, challenges and future directions," *Sensors*, 17(3) (2017): 1–42.
- [27] G. F. LOURENÇO, A. C. HONÓRIO, AND A. O. FIGUEIREDO, "Satisfaction of use of assistive devices for orientation and mobility of adults with visual impairment," *Journal of Occupational Therapy, University of São Paulo*, 28(3) (2017): 340–348.
- [28] W. ELMANNAI AND K. ELLEITHY, "Sensor-based assistive devices for visually-impaired people," *IEEE Access*, vol. 17(3), 2017, pp. 565–600.
- [29] B. RAJAPANDIAN, V. HARINI, D. RAKSHA, AND V. SANGEETHA, "A novel approach as an AID for blind, deaf and dumb people," *IEEE Access*, vol. 8, 2017, pp. 403–408.
- [30] S. MAIDENBAUM, S. ABOUD, AND A. AMEDI, "Sensory substitution: closing the gap between basic research and widespread practical visual rehabilitation," *Neuroscience & Biobehavioral Reviews*, 41 (2014): 3–15.
- [31] S. MAIDENBAUM, S. ABOUD, AND A. AMEDI, "Sensory substitution: closing the gap between basic research and widespread practical visual rehabilitation," *Neuroscience & Biobehavioral Reviews*, 41 (2014): 3–15.
- [32] A. FOLEY AND B. A. FERRI, "Technology for people, not disabilities: ensuring access and inclusion," *Journal of Research in Special Educational Needs*, 2012: 192–200.
- [33] D. DAKOPOULOS AND N. G. BOURBAKIS, "Wearable obstacle avoidance electronic travel aids for blind: a survey," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 40(1) (2010): 25–35.
- [34] V. TIPONUT ET AL., "Work directions and new results in electronic travel aids for blind and visually impaired people," *Proceedings of the 14th WSEAS International Conference on Systems*, Vol I, 2010, pp. 5346–5463.
- [35] A. IQBAL, U. FAROOQ, H. MAHMOOD, AND M. U. ASAD, "A low cost artificial vision system for visually impaired people," in *2009 Second International Conference on Computer and Electrical Engineering*, December 2009, pp. 474–479.
- [36] C. SHAH, M. BOUZIT, M. YOUSSEF AND L. VASQUEZ, "Evaluation of RU-Netra - Tactile Feedback Navigation System For The Visually Impaired," *2006 International Workshop on Virtual Rehabilitation*, 2006, pp. 72–77.
- [37] L. RAN, S. HELAL, AND S. MOORE, "Drishti: an integrated indoor/outdoor blind navigation system and service," in *Second IEEE Annual Conference on Pervasive Computing and Communications, 2004. Proceedings of the, IEEE*, 2004, pp. 23–30.
- [38] G. MOUNTAIN, "Using the evidence to develop quality assistive technology services," *International Journal of Integrated Care*, 2004: 19–26.
- [39] R. IMRIE AND P. HALL, *Inclusive Design: Designing and Developing Accessible Environments*, Spon Press, London, 2001.
- [40] M. J. SCHERER, *Living in the State of Stuck: How Technology Impacts the Lives of People with Disabilities*, Cambridge, MA, Brookline Books; ed 3, 2000.
- [41] MD. M. ISLAM, M. S. SADI, K. Z. ZAMLI, AND MD. M. AHMED, "Developing Walking Assistants for Visually Impaired People: A Review," *IEEE Sensors Journal*, vol. 19, no. 8, Jan 2019.
- [42] S. PONNADA, S. YARRAMALLE, AND M. R. T.V., "A Hybrid Approach for Identification of Manhole and Staircase to Assist Visually Challenged," *IEEE Access*, vol. 6, 2018.
- [43] N. S. AHMAD, N. L. BOON AND P. GOH, "Multi-sensor Obstacle Detection System via Model-based State-Feedback Control in Smart Cane Design for the Visually Challenged," *IEEE Access*, vol. 6, 29 Oct 2018, pp. 64182 - 64192.
- [44] J. BAI, S. LIAN, Z. LIU, K. WANG AND D. LIU, "Virtual-Blind-Road Following-Based Wearable Navigation Device for Blind People," *IEEE Transactions on Consumer Electronics*, vol. 64, no. 1, Feb. 2018, pp. 136–143.

*Edited by:* Anil Kumar Budati

*Special issue on:* Soft Computing and Artificial Intelligence for wire/wireless Human-Machine Interface

*Received:* Sep 22, 2023

*Accepted:* Dec 14, 2023



## OPTIMUM BATCH SCHEDULING MODEL FOR QUALITY AWARE DELAY SENSITIVE DATA TRANSMISSION OVER FOG ENABLED IOT NETWORK

NARAYANA POTU\*, CHANDRASHEKAR JATOTH† AND PREMCHAND PARVATANENI‡

**Abstract.** The emerging fog networks in the internet of things (IoT) applications provide flexibility and agility for service providers. The combination of fog nodes and edge nodes enable them to deliver a given network service. However, the selection of suitable edge and fog nodes and their scheduling still remain a research challenge. Finding a globally optimal scheduling of oversized data transmission over IoT applications for industrial requirements is crucial. Optimal batch scheduling has been regarded as a viable way to achieve optimal scheduling in other contemporary network models. This manuscript has projected an Optimum Batch Scheduling Model (OBSM) for Quality aware Delay Sensitive Data Transmission over Fog Enabled IoT Networks. A novel clustering technique has been proposed in this manuscript to group the transmission nodes (fog or edge nodes) and data packets, which further pairs each group of data with one of the corresponding node group to achieve delay sensitivity and other quality factors such as energy efficiency. The data scheduling between data and node group is drawn from the previous contribution -"Quality aware Energy Efficient Scheduling Model (QEESM) for Fog Enabled IoT Network". The simulation results have shown that, in terms of average make span rate, average round trip time, and energy consumption, the batch scheduling model OBSM performs noticeably better than the contemporary scheduling models. The OBSM scheduling model's average make-span rate, roundtrip time, as well as energy consumption per make span are 23.3 7.03, 17.8 5.2, and 11.33 6.9 joules, respectively, which conclusively demonstrate that the OBSM model outperforms the existing models. A novel batch scheduling algorithm has been proposed using a unique unsupervised learning approach that suggested to cluster the transmission requests and transmission channels in to multiple clusters.

**Key words:** Internet of Things, Industrial IoT, fog computing, edge nodes, Optimum Batch Scheduling Model, QALS

**1. Introduction.** Introduction. There is an increase in communication and computation latency and response times as a result of IoT devices generating a large volume, variety, and velocity of data [1]. For instance, IIoT gives connectivity among production lines and customers such that customers might guide the process of production directly as stated in [2], and connectivity of IIoT enables exchange of data among overall industrial devices such that entire production process could rapidly change by adapting to novel products [3]. The significant and fundamental part in IIoT is industrial networks.

In order to fulfill the crucial pre-requisites of industrial applications, the networks of industry have to transmit data, with accuracy in the control of data transmission, provision of sufficient bandwidth aimed at video-streams and handling each packet in huge communications. Several industrial wireless and wired networks have been projected for managing these high pre-requisites scenarios. Yet, there has been no one single network architecture, covering all the industrial pre-requisites. Heterogeneous networks comprising of wired network(s) and several wireless field networks are being considered as the future solutions for the industrial networks.

IoT has been a word for computing systems that refers to things created for rationally linking the animal world through the web, also known as the "things oriented vision," as mentioned in [4]. The architecture of the Internet of Things may be reduced to three mechanisms: devices, hardware, as well as middleware, depending on this previous contribution [5]. The first section of the system consists of the sensors and actuators needed for direct human and device communication in the physical world. Next, it is the responsibility of the gateways to enable and centralise communication among diverse objects. The third component, middleware, which is often found in the cloud, saves the data collected from other components as well as condenses the knowledge based on that data. The "REST (Representational state transfer)" web service is typically used by IoT middleware to adopt the SOA (service-oriented architecture).

---

\*University College of Engineering, Osmania University, Hyderabad, India (Corresponding author, [1potunarayana@gmail.com](mailto:1potunarayana@gmail.com))

†National Institute of Technology, Raipur, India ([chandrashekar.jatoth@gmail.com](mailto:chandrashekar.jatoth@gmail.com)).

‡University College of Engineering, Osmania University, Hyderabad, India.

The environment of IoT has been characterized by extreme heterogeneity where several devices of numerous architectures communicate among themselves along with other services of internet as in [5] [6]. Within this heterogeneous architecture, the fundamental role of gateways is twofold: they enable interaction of minimal computation performance systems to access internet providers as well as assure the secure communication of the wireless sensor nodes while the middleware interacts with diversified network protocols. However, they are faced with challenges: the gateways of IoT accept device data from several different communication networks; they need to retransmission the statistics to internet as well as convert the data received for Web service. The gateways would be further required to get packet data from hundreds of sensor systems as well as actuators. The tasks require a minimal latency in transmission as well as rapid data processing. In several instances, process bottle necks occur in REST call response time and not in transmission of packet. Thus, the usage of buffers has become essential for receiving and processing of data at the device end. In view of the extreme diversity of sensor data gathered by the devices, maximal transmission priority needs to be accorded for authentication of novel devices or actuators calls in the network. The gateway needs to prioritize the urgent calls over regular ones in the incoming traffic. Thus, the gateway will have to address the problems related to huge amount of data volumes, prioritization, processing of packets, synchronous messaging, and performance difference among network mechanisms.

There could be an effect when the performance of various algorithms and protocols changes. Some of the sensor devices, for example, use TCP transport, while others use UDP, and still others don't have a transport layer at all. It is also possible to mention connection and physical-layer strategies, such as the (IEEE 802.11) Wi-Fi network or the (IEEE 802.15.4) ZigBee network. Furthermore, when the gateway uses HTTPS protocol, sensors often use device manufacturer's proprietary protocols in the layer of application. In addition, this cross-pollination of technologies adds IoT gateway's computational complexity. The packet's priority is to recognize and manage packets with the highest priority. For instance, sensors that has been evaluating something of extreme importance, sensor components that have demanded action through actuators, or even tasks of high priority, such as sensor authentication in a network before forwarding data packets. In the case of Synchronous Messaging, synchronous links between middleware and sensor devices may be a bottleneck because the middleware's for the following packet to be processed, a response is necessary. Additionally, these difficulties call for the application of IoT gateway methods including traffic shaping, packet discarding, and queue management. In order to provide effective and equitable treatment again for requests processed by the gateway, this study proposes an IoT gateway including QoS qualities based on a network prioritization algorithm.

**1.1. Motivation.** The evolving IoT is thought of as the next internet generation. Actuators, cars, phones, cars, and sensors are examples of things that interact to provide a service. Cloud computing is a new computing model that allows users to access a shared pool of resources on demand. The cloud and IoT are currently attracting both industry and academia's attention. Sensor data is sent to a gateway, which sends it to the cloud. Typically, the cloud stores, processes, transmit, and stores user-generated data. When data transmission from sensor nodes to the cloud fails, data is re-transmitted until it succeeds.

**1.2. Problem statement.** In the fog IoT prototype, users serve as data requesters and the IoT serves as a source of data for the cloud. When there is an internet access, users may be able to get the necessary sensor data from the cloud. All of these applications for cloud-based IoT integration, such as smart cities and buildings depend on IoT to consistently deliver sensory data to the cloud in response to user requests [7]. Since non-rechargeable batteries are used in most sensors, data sensing performance, transmission, and processing depletes battery power over time. Several models exist for optimizing power consumption and improving IoT reliability. However, power consumption reduction models have negatively impacted network reliability.

**1.3. Organizing of the Manuscript.** In the further sections of this manuscript, section-2 refers to the related work summary from the literature review. Section-3 provides insights into the materials and methods, proposed model narrative, its algorithm flow, and other key metrics that signify the mode. Section-4 provides insights into experimental study and section-5 refers to the conclusion based on the efficiency aspects estimated from the model.

**2. Related Research.** According to Liu et al., [8], an ADEC (adaptive dynamic energy consumption) improvement strategy was proposed based on the JNCC model. The Improved Software-Defined WSN was developed by Y Duan et al., [9]. (Improved SD-WSN). Managing the network and ensuring adequate coverage are the primary concerns. To reduce energy consumption, C Zhu et al., [10] proposed a novel WSN-mobile cloud computing integration scheme that includes two parts: 1) a prioritized sleep scheduling algorithm for WSN and 2) a time-based selective data transmission algorithm for mobile cloud computing.

Each sensor's computing load and energy consumption are both reduced thanks to their use of fog-cloud hierarchical network architecture. Wang et al., [11] that a single, energy-constrained source can be used to power multiple energy-constrained relays in a wirelessly powered Internet of Things by selecting the best power beacons (PBs) (IoT), have proposed it. Over the Rayleigh fading channel, each scheme has closed-form expressions for power outage, secrecy outage, and energy efficiency IoT-cloud service prototype by S Kim et al., used Docker Swarm-based container orchestration to ensure the service's reliability and tested its uptime [12].

Each application's delay bound and target reliability were guaranteed using a QoS framework for arbitrary hybrid wired/wireless networks by S Zoppi et al., [13]. An alternative scheduling algorithm for wireless sensor networks is also proposed. Three methods are recommended by Y Peng and colleagues [14]: short, ongoing, and noise-related faults. Using an algorithm that estimates probability-guaranteed limits on packet reception ratio, W Sun et al., [15] propose a smart grid solution. According to J Yuan et al., [16], trusting IoT edge devices requires a multi-source feedback information fusion approach. For large-scale IoT edge computing, they present a low-overhead trust evaluation mechanism. Traditional trust schemes can be replaced by a new algorithm that incorporates feedback from multiple sources. In order to allow any TSCH node to transmit or receive frames at any time, H Park et al., [17] and his colleagues developed STATIC TSCH scheduling. Large-scale smart meter networks can be built with TSCH. Broadcast and unicast slots were defined separately to prevent network control message collisions.

In order to create an IoT platform with high exile and reliability, S C Wang et al., [18] used fog and cloud computing in conjunction (IFCIoT). IFCIoT can be used in a variety of other applications and disaster monitoring systems. All fault-free nodes in the IFC IoT platform can come to an agreement on a protocol with the smallest number of messages, while still allowing for malicious and dormant components. D Purkovic et al., [19] have proposed an energy-efficient communication protocol. In order to collect environmental data with the least amount of energy possible, this protocol is used. It's designed for sensors with short battery lives and energy harvesting. Teach-in and Data Telegram packets are used to collect the data.

JuanLuo et. al [20] proposed to increase the system efficiency of fog devices as well as decrease service latency, a multi-cloud into multi-fog architecture is used, along with the construction of two different service models using containers. The MILP was used by Al-Shammari et al., [21] to investigate the energy efficiency of a smart city service embedding framework. Using this framework, IoT systems will be able to meet the virtual node and link requirements of business processes while using less power. Fog computing for health monitoring was investigated by Ida Syafiza M. Isa et al., [22]. They used a MILP model to process and analyze patient electrocardiogram signals. Use of the most energy-efficient locations for processing and networking servers.

PegahGazori et. al. [23] concentrated on scheduling tasks in fog-based IoT systems with the goal of minimizing length of service latency and computation cost while adhering to resource and deadline limitations. We have conducted research using the reinforcement learning technique to solve this issue [23]. An IoT-based real-time HVAC control system [24] was designed and implemented using thermal comfort, demand response, and user feedback. In order to forecast the room's thermal parameters, they use artificial neural networks trained on historical data. With the help of MILP, the HVAC control problem is optimized for energy efficiency and user satisfaction. Decentralized micro grid energy exchange mechanisms have been proposed by Laszka et al., [25]. Consumers no longer have to worry about their privacy or the security of the system when they trade energy. In order to quickly and securely clear offers, the platform uses a hybrid MILP solver approach. IoT service selection was proposed by M E Khanouche et al., [26]. Energy consumption is minimized while service quality requirements are met by this multi-objective optimization problem. Pre-selecting services that meet the quality of service (QoS) requirements of end-users is a goal of each others. Pareto's concept of relative dominance relations is used to select the best service. It is the user's preferences that determine the relative dominance of a potential service (QoS). Near-optimal solutions for large-scale IoT environments with thousands

of distributed entities can be found using the EQSA algorithm (about 98 percent).

The Contemporary models “QEESM (Quality Aware Energy Efficient Scheduling Model) [27]”, “CFCA (Container based Fog Computing Architecture) [28]”, and “QALS (Quantum approach of load scheduling) [29]” are competent methods to perform data transmission with minimal droop ratio across IoT networks using fog computing. However, none of the aforesaid methods considering the context of the oversized delay sensitive data transmission. Considering the scope of a scheduling model to perform optimal transmission of oversized delay sensitive data transmission over fog enabled IoT networks, this manuscript endeavoured to portray an Optimum Batch Scheduling model for Quality aware Delay Sensitive Data Transmission over Fog Enabled IoT Networks.

**3. Methods and Materials.** The optimum batch scheduling model have been addressed to achieve delay sensitive and quality aware data scheduling In IoT networks with fog support. The issue of data frame loss is critical in IoT networks, which is due to overloaded transmission or by any crux of network quality issues that often admitted by IoT devices with weak resources. Hence, suboptimal data scheduling in IoT networks is a significant research issue to address. In order to this, earlier contribution has suggested a novel scheduling strategy for Fog enabled IoT networks that intended to achieve energy efficiency and other quality factors [27]. However, the overwhelmed data transmission sources of IoT networks, which has been connected to fog network based cloud services such as industrial IoT networks still causes considerable data frame drop ratio, which often critical to address in delay sensitive fog enabled IoT networks. To address this issue, this contribution has portrayed a batch scheduling strategy, which is a firm extension of the earlier contribution “Quality-aware Energy Efficient Scheduling Model (QEESM) for Fog Enabled IoT Networks”. The suggested batch scheduling strategy is an unsupervised clustering technique that partitions the data frames to be transmitted over fog enabled IoT networks in to multiple clusters and then performs QEESM scheduling model on each cluster in the priority order of these clusters defined under delay sensitivity. The subsequent sections explore the methods and materials used in scheduling process portrayed [30], [31].

**3.1. Data-frames and node clustering process.** The process of clustering the nodes is initiated based on the estimation of distance amidst the idle times among the fitness function. Every cluster signifies the nodes based on the idle time overlap conditions. The fundamental function in the process of nodes clustering is to focus on the start and end time related to the nodes to observe overlap conditions. Accordingly, all the overlapping nodes are grouped into one. In line with the above the mentioned process, even the data-frames are grouped in order that the data-frames overlap based on the time-interval feasible for handling arrival and residual session span.

The process adapted in the case of the data-frames streaming to scheduler on basis of incremental arrival time, and the time interval nodes. The process refers to the conditions wherein the incremental start-time is considered, as well as the cluster - based method is triggered as a positive approach, to mitigate the risk of data-frame loss or inappropriate scheduling of the nodes. The objective is to overcome certain challenges of clustering which are imperative in the existing solutions.

In the proposed solution, the clustering process adapted is inspired by the traditional k-means clustering algorithm [32], [33]. However, the change considered in the model is about unrestricting the cluster counts. The emphasis is more about adhering to time series compatibility, wherein the newly arrived record relates to the recently formed cluster or shall be allowed to form a new cluster. The new cluster formation shall be based on the steps discussed in the sub-sections.

**3.1.1. Node clustering.** .When no more nodes satisfy the requirements, a cluster is considered complete. The process proceeds with the remaining nodes after a cluster’s nodes are eliminated from the list. Until all nodes are clustered, this process is repeated. Idle nodes are effectively grouped using node clustering algorithms. Using a pre-sorted list, the methodical approach clusters nodes based on overlaps in idle time. Figure 3.1 illustrates how this algorithm forms clusters gradually to maximize network performance. Nodes are arranged to maximize idle times using this thorough clustering technique, which enhances network performance.

#### **Algorithm to cluster the idle nodes**

The objective of the algorithm is to cluster nodes based on their idle time intervals, optimizing network

performance in a fog-enabled IoT environment. The algorithm operates on a set of nodes with defined idle time intervals and employs a clustering mechanism inspired by k-means but adapted for time interval-based clustering.

1. **Initialization and Node Sorting:** Let  $L = \{e_1, e_2, \dots, e_n\}$  represent the list of nodes, sorted in ascending order based on the start times of their idle time intervals. This ensures  $idle\_start(e_i) \leq idle\_start(e_{i+1})$  for all  $i$ .
2. **Cluster Centroid Initialization:** The first node in the list,  $e_1$ , is initialized as the centroid for the first cluster,  $ctrd_j$ , of the  $j^{th}$  cluster  $ncl_j$ .
3. **Cluster Formation:**
  - For each node  $e_i$  in  $nL$ , the algorithm checks if  $idle\_start(e_i) < idle\_start(ctrd_j) + idle\_duration(ctrd_j)$ .
  - If the condition is satisfied,  $e_i$  is added to  $ncl_j$ .
4. **Updating Clusters and Centroids:**
  - If new nodes are added to  $ncl_j$ , the algorithm sorts the nodes within the cluster in decreasing order of their idle time durations.
  - The node with the longest idle duration in  $ncl_j$  becomes the new centroid,  $ctrd_j$ .
  - If no new nodes are added, indicating that the cluster is complete,  $ncl_j$  is finalized.
5. **Iteration:**
  - Remove the nodes of  $ncl_j$  from  $nL$  and increment  $j$  for the next cluster formation.
  - Repeat the process until all nodes in  $nL$  are clustered.
6. **Termination:**
  - The process ends when  $nL$  is empty, indicating all nodes have been clustered.

This algorithm effectively groups nodes based on overlapping idle time intervals, optimizing the utilization of resources in a fog computing network. The mathematical approach ensures precision in clustering, leading to enhanced network performance and efficiency.

**3.1.2. Data-frames clustering process for schedule..** The scheduling process for idle frame interval nodes is similar to that of data-frame clustering. Data frames are buffered and grouped by crucial transmission times by the scheduler shown in figure 3.2. In this process, data frames are arranged according to arrival times. The first data-frame in this ordered list, denoted as  $df$ , serves as the centroid of the first cluster, denoted by  $dfcl_1$ . Making this decision creates the foundation for cluster formation. The data-frame arrival times and centroid transmission times are compared using the clustering method. Data frames whose arrival times coincide with the centroid's transmission interval are received by the cluster  $dfcl_1$ . As a result, data frames with comparable transmission properties are grouped. During the process, the data-frame from  $dfcl_1$  with the maximum transmission end time is used to reform the cluster centroid. To capture the most representative transmission characteristics of the cluster, dynamic centroid selection is essential. When the composition of a cluster remains constant between centroid formations, it is considered final. Following finalization, these clusters are eliminated from the list, and the remaining data frames are processed in the same manner. Until every data frame is grouped into clusters, this cycle is repeated. The interval start list is replaced by the data-frame list (DFL) and the arrival list of each data-frame, denoted as  $a(df)$ , which are represented by the evolutionary clustering algorithm for this frame. The transmission time that is needed is  $rtt(df)$ . This cloning of the clustering process is depicted in the algorithmic representation below.

#### Algorithm: Data-Frames Clustering

1. **Initialization:**
  - Define  $dfl = \{df_1, df_2, \dots, df_n\}$  as the list of data-frames, sorted in ascending order based on their arrival times,  $a(df_i)$ .
  - Define  $df_i^{rt}$  as the required transmission time of data-frame  $f_i$ .
  - Initialize the cluster set,  $dfcl_j = \phi$ , and a temporary set,  $ts = \phi$ .
  - Set the initial cluster centroid,  $trdf_j$ , as the first data-frame in  $dfl$ .
2. **Clustering Process:**
  - For each data-frame  $df_i$  in  $dfl$ :

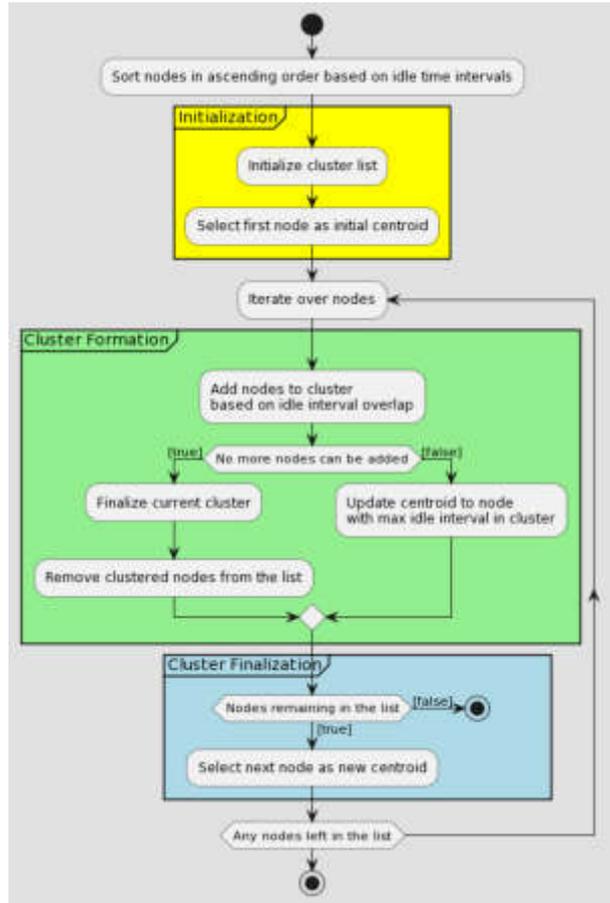


Fig. 3.1: FlowChart Representing of the OBSM Node Clustering

– If  $a(df_i) < a(ctrdf_j) + df_j^{ct}$ , add  $df_i$  to  $dfcl_j$ .

**3. Centroid Update and Cluster Finalization:**

- If new data-frames are added to  $fcl_j$ , determine the data-frame with the maximum sum of arrival time and required transmission time, denoted as  $mrss$ .
- Update  $ctrdf_j$  to the data-frame with the maximum  $mrss$ .
- If no new data-frames are added, finalize  $dfcl_j$  and remove its elements from  $dfL$ .

**4. Iteration and New Cluster Formation:**

- If  $dfL$  is not empty, increment the cluster index  $j$  and repeat the process with the remaining data-frames in  $dfL$ .
- Set  $ctrdf_j$  as the first data-frame in the updated  $dfL$ , and initialize a new cluster  $dfcl_j$ .

**5. Termination:**

- The process terminates when  $dfL$  is empty, indicating all data-frames have been clustered.

**3.1.3. Ranking of the data-frame and cluster of nodes.** . Each cluster of idle nodes is correlated with the data-frames using specific criteria, allowing the process to estimate three objectives. Distances are defined as the minimum start time of idle intervals for nodes within clusters and the minimum arrival time of data-frames. A comparison is also made between the data-frame transmission time and the mean idle time in clusters. Additionally, the process makes a comparison between the mean intervals of idle transmission times for nodes in the same cluster and the variance in data frame transmission times for nodes. Idle node clusters

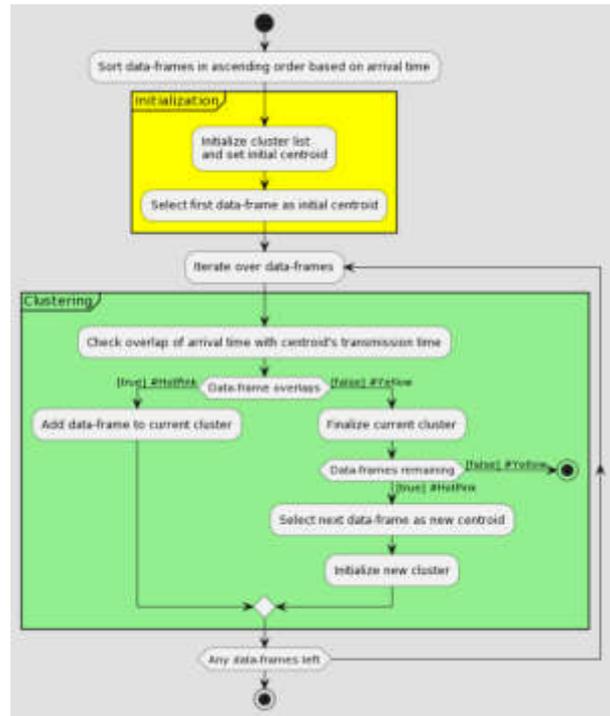


Fig. 3.2: Flow Chart Representing of the OBSM Data Frame Clustering

can be identified by classifying idle clusters by least arrival time data in descending order. Based on their sorted list position, these clusters are ranked according to arrival rate. The next step is to rank the data-frame clusters according to position and sort them according to mean transmission time. Based on average idle time intervals, clusters are ranked. Data-frame clusters are sorted in descending order using the variance in transmission times. Based on the variance of idle intervals, idle node clusters are ranked. The intervals between data transmission ranks, idle times, and transmission time deviation ranks are all estimated by this process. For every data-frame cluster, a notable cluster of nodes is selected in order to accomplish several objectives. The algorithmic ranking process for data-frame clusters is presented below. Please only use this process for Data-Frame Clusters.

The set of data-frame clusters, DFCL, is represented as  $dfcl_1, dfcl_2, \dots, dfcl_{|DFCL|}$ . The algorithm involves cloning DFCL into a temporary set TC and creating ordered sets for data-frame clusters based on the least arrival time, average residual session span, and deviation in transmission times. The algorithm iterates through these sets, identifying data-frame clusters with the least arrival time, session span, and deviation. For each identified cluster, a rank is assigned based on its order in terms of arrival time, transmission time, and deviation. These ranks are used to move the data-frame cluster into respective ordered sets. The process continues until all clusters in TC are ranked.

#### Algorithm signifying discriminative ranks for data-frame clusters

##### 1. Initialization:

- Let  $DFCL = \{dfcl_1, dfcl_2, \dots, dfcl_n\}$  be the set of all data-frame clusters.
- Define  $aoDFCL$ ,  $soDFCL$ , and  $doDFCL$  as sets for arranging data-frame clusters based on least arrival time, average session span, and deviation in transmission time, respectively.
- Initialize an index counter  $idx$  to 0.

##### 2. Ranking Based on Arrival Time:

- For each cluster  $dfcl_i$  in DFCL:
  - If  $dfcl_i$  is not in  $aoDFCL$ , find the cluster with the least arrival time and assign it to  $dfcl_{ia}$ .

- Update  $aoDFCL$  by adding  $dfcl_{ia}$  and assign it a rank based on its position in the ordered set.
3. **Ranking Based on Session Span:**
    - For each cluster  $dfcl_i$  in DFCL:
      - If  $dfcl_i$  is not in  $soDFCL$ , find the cluster with the least session span and assign it to  $dfcl_{is}$ .
      - Update  $soDFCL$  by adding  $dfcl_{is}$  and assign it a rank based on its position in the ordered set.
  4. **Ranking Based on Deviation:**
    - For each cluster  $dfcl_i$  in DFCL:
      - If  $dfcl_i$  is not in  $doDFCL$ , find the cluster with the least deviation and assign it to  $dfcl_{id}$ .
      - Update  $doDFCL$  by adding  $dfcl_{id}$  and assign it a rank based on its position in the ordered set.
  5. **Increment Index and Repeat:**
    - Increment the index  $idx$ .
    - Repeat the ranking process for each cluster in DFCL until all clusters have been ranked in  $aoDFCL$ ,  $soDFCL$ , and  $doDFCL$ .
  6. **Output:**
    - The output of the algorithm is the sets  $aoDFCL$ ,  $soDFCL$ , and  $doDFCL$  with their respective ranks, representing the discriminative ranks of data-frame clusters based on arrival time, session span, and deviation in transmission time.

**3.1.4. Collation of data-frame clusters and clusters of nodes.** . The correlation of data-frame clusters and node clusters is based on the discriminative ranks allocated in the earlier process. For a specific data-frame cluster  $dfcl$ , a corresponding cluster of nodes  $ncl$  is determined based on several rank comparisons. These include the comparison of the rank of idle intervals in  $ncl$  with the mean transmission time rank of  $dfcl$ , the variance span rank in  $ncl$  with the variance span rank of  $dfcl$ , and the start time of idle intervals in  $ncl$  with the least arrival time rank of  $dfcl$ . Scheduling of data-frames is managed according to these paired clusters of nodes, ensuring an optimized correlation between node availability and data-frame transmission requirements.

**3.2. Residual energy.** The residual-energy pertaining to the  $n^{th}$  make-span at  $k^{th}$  transmission node is considerably the remains of energy observed from the earlier make-span ( $(n-1)^{th}$  completion). Residual-energy  $re_n$  for the  $n^{th}$  make span is estimated as

$$re_n = (re_{(n-1)} + ce_{(n-1)} - coe_{(n-1)}) - cot$$

The notations  $re_{(n-1)}$ ,  $re_n$ ,  $ce_{(n-1)}$ ,  $cot$  are used for indicating the residual-energy for the  $(n-1)^{th}$ ,  $n^{th}$  make-spans, wherein energy conserved for the period as the  $(n-1)^{th}$  make-span, and the energy consumed at  $(n-1)^{th}$  make-span and the energy consumed in the process of idle time amidst  $(n-1)^{th}$  and  $n^{th}$  make-spans of the target node.

To assess the optimality for the projected  $(n+1)^{th}$  the make-span, transmission node process constitutes energy efficiency, turnaround time ratio, and Task Arrival Time Interval ratio. The scheduling strategy proposed for handling the process is as follows. For every metric, turnaround time interval, the scope of process completion time interval, turnaround time interval, and task arrival interval are estimated for the make-span  $(n+1)^{th}$ , wherein the Max-State ( $mas$ ), Min-State ( $mis$ ), and Close-State ( $cls$ ) are observed in accordance to how the make-span close.

In furtherance, a two-dimensional matrix for every metric is depicted wherein the values for the status measures are handled in a matrix format depicted below Table 3.1.

Followed by, for every metric, the task arrival time interval, turn around time interval, and process completion time interval are handled effectively wherein the moving averages for every status measure are assessed based on the close-state, max-state, min-state, and initial-state. For every column  $c$  related to the two-dimensional matrix  $M(ins)$ ,  $M(mas)$ ,  $M(mis)$ , or  $M(cls)$  referring to the status measures  $ins$ ,  $mas$ ,  $mis$ , and  $cls$  of every metric as follows: turnaround time interval ( $tti$ ), task arrival time interval ( $tati$ ), process com-

Table 3.1: The values for the status measures are handled in a matrix

Make span-ID	Initial-State	Max-State	Min-State	Close-State
1	$[ins_1^{mi}]^{-1}$	$[mas_1^{mi}]^{-1}$	$[mis_1^{mi}]^{-1}$	$[cls_1^{mi}]^{-1}$
2	$[ins_2^{mi}]^{-1}$	$[mas_2^{mi}]^{-1}$	$[mis_2^{mi}]^{-1}$	$[cls_2^{mi}]^{-1}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
i	$[ins_i^{mi}]^{-1}$	$[mas_i^{mi}]^{-1}$	$[mis_i^{mi}]^{-1}$	$[cls_i^{mi}]^{-1}$
$i + 1$	$[ins_{(i+1)}^{mi}]^{-1}$	$[mas_{(i+1)}^{mi}]^{-1}$	$[mis_{(i+1)}^{mi}]^{-1}$	$[cls_{(i+1)}^{mi}]^{-1}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
n	$[ins_n^{mi}]^{-1}$	$[mas_n^{mi}]^{-1}$	$[mis_n^{mi}]^{-1}$	$[cls_n^{mi}]^{-1}$

pletion time interval (pcti), The necessary moving averages for each of the column are observed, for column  $c$ .

$$ma_i(c) = \frac{1}{mac} \left( \sum_{j=i}^{i+mac-1} c[j] \right), \quad \text{for } i = 1 \text{ to } (|c| - mac) \tag{3.1}$$

Eq 3.1 moving average referring to values of column  $c$  of the two-dimension matrix  $M$  of the metric  $tti$ ,  $tati$ , or  $pcti$  Post the process, for every set of moving average, the Heikin-ashi equivalent is developed based on the open, low, high, as well as end values, which enables in deriving the candle patterns for representing the expected make-suitability span's for the associated metric.

Depending on the indicated nodes, the emphasis is optimality for all the metrics like the Heikin-Ashi candle patterns, which are estimated. For each node  $n$ , for every metric  $rt, rr, cl$ , the moving average with each  $maM(rt), maM(rr), maM(cl)$ , The following is how Heiken-ashi patterns are created.  $ha - maM_i(o) = maM_i(o) //$  For the moving averages for the initial make-span-related indicators, Else, the Heiken-ashi moving averages are calculated as follows: Eq 3.2 to Eq 3.5

$$ha - maM_i(o) = \frac{ha - maM_{(i-1)}(o) + ha - maM_{(i-1)}(e)}{2} \tag{3.2}$$

$$ha - maM(e) = \frac{maM(o) + maM(h) + maM(l) + maM(E)}{4} \tag{3.3}$$

$$ha - maM(l) = \min(ha - maM(o), maM(l), ha - maM(e)) \tag{3.4}$$

$$ha - maM(h) = \max(ha - maM(o), maM(h), ha - maM(e)) \tag{3.5}$$

Subsequently, it determines each node's grade coefficient in the manner shown below.

For every node  $n$ , for every metric a request receiving time ( $rr$ ), round-trip time ( $rt$ ), as well as computational-process-time ( $cl$ ) are used for observing the average of the high-value ( $h$ ), status-measures (open-value ( $o$ ), end-value ( $e$ ), as well as low-value ( $l$ ), as well as a variance, the root-mean-square length using the equation as: Eq 3.6 to Eq 3.14.

$$\langle rt_n \rangle = \frac{rt_n^o + rt_n^h + rt_n^l + rt_n^e}{4} \tag{3.6}$$

$$rt_n^p = \frac{\sqrt{\langle rt_n \rangle - rt_n^o{}^2} + \sqrt{\langle rt_n \rangle - rt_n^h{}^2} + \sqrt{\langle rt_n \rangle - rt_n^l{}^2} + \sqrt{\langle rt_n \rangle - rt_n^e{}^2}}{4} \tag{3.7}$$

Table 3.2: Optimal Node IDs

Node ID	Quality Coefficient	Projected Energy Consumption	Residual Energy
---------	---------------------	------------------------------	-----------------

Table 4.1: The Table Record ID, Metric Values

RECORD-ID	METRIC VALUES			
Node-ID	High-value	Open-Value	End-value	Low-value
Makespan-ID				
Metric-ID				

$$rt_n^{mc} = \langle rt_n \rangle + rt_n^p \quad (3.8)$$

$$\langle rr_n \rangle = \frac{rr_n^o + rr_n^h + rr_n^l + rr_n^e}{4} \quad (3.9)$$

$$rr_n^p = \frac{\sqrt{(\langle rr_n \rangle - rr_n^o)^2} + \sqrt{(\langle rr_n \rangle - rr_n^h)^2} + \sqrt{(\langle rr_n \rangle - rr_n^l)^2} + \sqrt{(\langle rr_n \rangle - rr_n^e)^2}}{4} \quad (3.10)$$

$$rr_n^{mc} = \langle rr_n \rangle + rr_n^p \quad (3.11)$$

$$\langle cl_n \rangle = \frac{cl_n^o + cl_n^h + cl_n^l + cl_n^e}{4} \quad (3.12)$$

$$cl_n^p = \frac{\sqrt{(\langle cl_n \rangle - cl_n^o)^2} + \sqrt{(\langle cl_n \rangle - cl_n^h)^2} + \sqrt{(\langle cl_n \rangle - cl_n^l)^2} + \sqrt{(\langle cl_n \rangle - cl_n^e)^2}}{4} \quad (3.13)$$

$$cl_n^{mc} = \langle cl_n \rangle + cl_n^p \quad (3.14)$$

In furtherance, sorts of the optimal nodes  $oN$  as an ascending order for the projected residual energy are estimated, as it can refer to the node scope in terms of completing the transaction, and sorting in open-value, feasible end-value. Also, in terms of the descending order for the respective quality coefficient, the nodes sort shall appear in the following dimension Table 3.2.

Thus, the scheduling strategy chosen as one of the significant issues in the listed nodes constitute to schedule based on the contextual factors considered, as necessary.

**4. Experimental Study.** The suggested Optimum Batch Scheduling Model (OBSM) for quality aware delay sensitive data transmission across fog enabled IoT networks has undergone an experimental evaluation to scale performance.

The experimental study performed in passive model, which executes the proposed and contemporary models QEESM, on known data (labelled data). The dataset has generated and used, which explored in contemporary model QEESM. The dataset format has been portrayed following Table 4.1.

This aforementioned dataset was created using a simulation based on fog-sim. The following table lists the simulation parameters that were employed (Table 4.2).

A total of 3600 internet protocol equipped sensors have been taken into account when creating the dataset. Each of them produces data at an average speed of 300 kbps. Through the scheduling gateway, these sensors are connected to the edge nodes. Additionally, these edge nodes are connected to a network of 50 fog nodes in a fog

Table 4.2: The Simulation Parameters

Parameter	Specification
Low power parameter	eDRX, Power saving mode
Latency parameter	greater than 10 seconds
Data Rate specification	25 kbps in download and 64 kbps in UL
Link budget parameter	above 164 dB (20dB GPRS)
Modulation scheme	
Uplink parameter	$\pi/4$ -QPSK, $\pi/2$ -BPSK, QPSK
Downlink parameter	QPSK
Multiple access	
Downlink specification	OFDMA
Uplink specification	SC-FDMA
Duplex Mode specification	FDD Half Duplex Type B
Frequency Range parameters	1,2,3,5,8,11,12,13,17,18,19,20,25,26,28,66,70 MHz
Supporting parameters	Uplink power control, HARQ

Table 4.3: Average Make-span Rate for Both Current Models and OBSM

Node Count	10	15	20	25	30	35	40	45	50
OBSM	13	15	15	20	28	30	29	29	31
QEESM	17	19	19	23	31	32	32	33	34
CFCA	19	22	22	26	35	36	35	36	37
QALS	35	37	41	44	46	49	51	54	55

computing system. The metric data collected from the most recent 60 make-spans in order are displayed one per node of the edge and fog networks. The dataset that was produced has 216,000 records in total. It's make-span rate on average, round-trip duration beside a varied set of fog devices, as well as variable transmission load are the metrics taken into account for performance analysis. The energy consumption ratio versus fluctuating load is another key metric that was evaluated.

By making comparisons the metric reported the results for OBSM with said correlating measure obtained values again from recent methods QEESM [27], CFCA [28], as well as (QALS) [29], it can be seen that the effectiveness of the OBSM has increased.

The presence of the nodes is indicated by the metric 'make-span rate,' which is proportional to each other. The lower make-span rate is predicted by Table 4.3 and Figure 4.1 to be inversely proportional to the total fog nodes. The typical make-span rate that the OBSM scheduling model has detected is  $23.3 \pm 7.03$ . When compared to the mean make-span rates  $26.7 \pm 6.62$ ,  $29.8 \pm 6.95$ , and  $45.8 \pm 6.7$  of the QEESM, CFCA, and QALS models, the mean make-span rate of the OBSM is noticeably low and has a little divergence.

For OBSM, QEESM, CFCA, QALS, other reliability of this research round frame has indeed been taken into account as well as compared (see Table 4.4, Figure 4.2), The statistics shown in Figure 4.2 clearly show that the model OBSM outperforms the current models QEESM, CFCA, QALS with shorter round-trip time. The mean of the average round-trip times recorded from OBSM, QEESM, CFCA, QALS is  $17.8 \pm 5.2$ ,  $33.96 \pm 6.0$ ,  $40.36 \pm 10.3$ , as well as  $44.56 \pm 14.3$ .

The anticipated round-trip time for a variable load is shown in Table 4.5 and Figure 4.3. When scaled against the modern models QEESM, CFCA, QALS, the model OBSM has shown a superior performance. The average difference here between round-trip time  $14.9 \pm 7.7$  of the OBSM, and the round-trip time  $32.3 \pm 7.01$  of the QEESM,  $39.9 \pm 12.9$  of CFCA, and  $44.9 \pm 17.9$  of QALS model.

Table 4.6 and Figure 4.4 show statistics for the essential objective energy consumption. The mean energy consumption in joules for every make-span for the proposed as well as current models, QEESM, OBSM, CFCA,

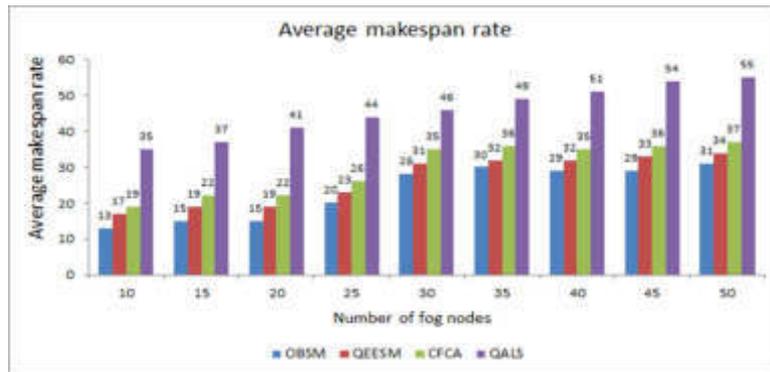


Fig. 4.1: The common make-span rates for the OBSM, QEESM, CFCA, and QALS

Table 4.4: Average Round-Trip Time for Fog Nodes with a Changeable Number of Nodes vs. a Fixed Load (Specify Constant Load Value)

Node Count	10	15	20	25	30	35	40	45	50
OBSM	21	25	22	21	19	18	15	11	8
QEESM	42	40	39	37	33	31	29	28	23
CFCA	44	42	43	40	36	35	33	33	27
QALS	49	46	47	44	40	38	36	36	30

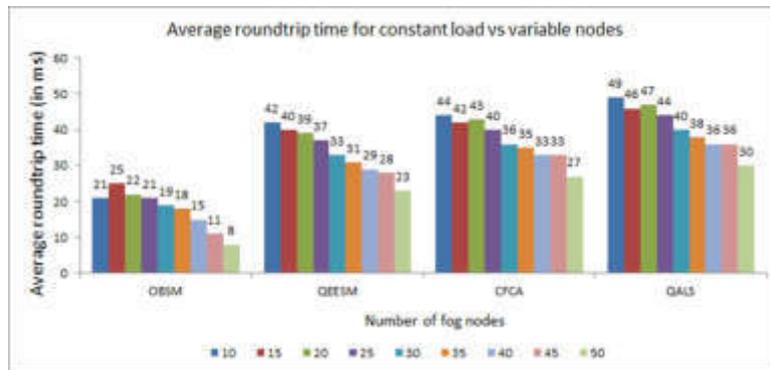


Fig. 4.2: The average round-trip time of OBSM, QEESM, CFCA, and QALS versus variable number of fog nodes

Table 4.5: Average Round-Trip Time for a Load Fixed vs a Different Number of the Fog Nodes

Load in KBPS	100	150	200	250	300	350	400	450	500
OBSM	5	8	9	14	14	14	16	22	32
QEESM	22	25	28	30	29	34	38	41	44
CFCA	28	30	32	34	35	40	45	46	51
QALS	33	35	37	38	40	44	48	52	55

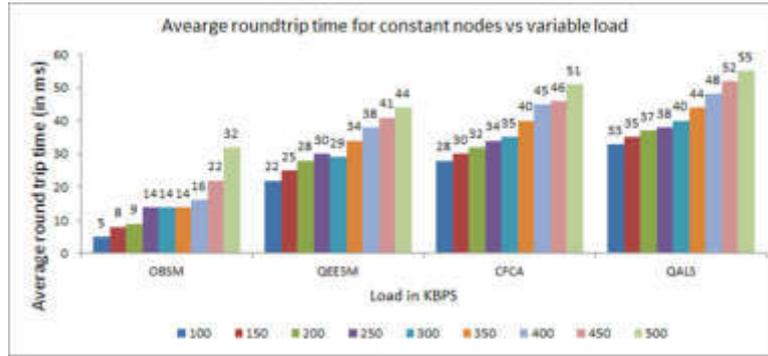


Fig. 4.3: Models OBSM, QEESM, CFCA, & QALS round-trip time statistics versus variable nodes

Table 4.6: Energy consumption for each make-span with a variable load

Load in KBPS	100	150	200	250	300	350	400	450	500
OBSM	1	2	6	9	12	15	17	18	22
QEESM	3	9	12	20	21	25	31	39	44
CFCA	8	14	17	24	26	29	35	44	49
QALS	12	18	21	29	31	35	39	48	53

as well as QALS, has been evaluated by comparing. The results show that the proposed model, including an average of  $11.33 \pm 6.9$  joules for every make-span, outperforms a current model, QEESM with a consumption of  $22.7 \pm 12.8$  joules per make-span, CFCA with a consumption of  $27.4 \pm 12.8$ , and QALS. In comparison to the energy used by the modern models QEESM, CFCA, and QALS, the OBSM uses less energy on average.

**4.1. Critical Analysis.** An experimental study on the Optimal Batch Scheduling Model (OBSM) assesses how well it manages delay-sensitive, quality-aware data transmission over fog-enabled Internet of Things networks. In this study, a specially generated dataset and an elaborate simulation setup are used to compare the OBSM with contemporary models such as QEESM, CFCA, and QALS. A widely used IoT simulation tool called fog-sim was used to create the dataset for this study. High, open, end, and low metrics for a thorough analysis are displayed in Table 4.1.

The simulation parameters in Table 4.2 are essential to the experimental setup used in this study. Examples include data rate specifications, latency parameters longer than 10 seconds, and low power parameters such as eDRX and Power Saving Mode. The modulation scheme, downlink and uplink parameters, and auxiliary features like HARQ and Uplink Power Control are all included in the simulation. The experiment's applicability to actual Internet of Things network scenarios is enhanced by meticulous parameter selection.

3600 internet protocol-equipped sensors are considered in the experiment as part of a network setup. A fog computing system is created by connecting these sensors through edge nodes to a network of fifty fog nodes. A critical analysis is conducted on the make-span rate, round-trip duration, energy consumption ratio, and over variable fog device and transmission load counts. An extensive setup like this offers a scalable and realistic OBSM testbed.

A thorough comparative analysis of OBSM is given in Tables 4.3 through 4.6 and Figures 4.1 through 4.4. A crucial network node indicator, the make-span rate, is displayed in Table 4.3 and Figure 4.1. The OBSM exhibits a lower make-span rate across node counts, suggesting improved task handling. Figures 4.2 and 4.3 and Tables 4.4 and 4.5's round-trip time comparisons, which show that OBSM reduces transmission delays more quickly, corroborate this.

The analysis of energy consumption (Table 4.6, Figure 4.4) is an important aspect of the study, showing that OBSM is more energy-efficient than the other models. This finding highlights the applicability of the

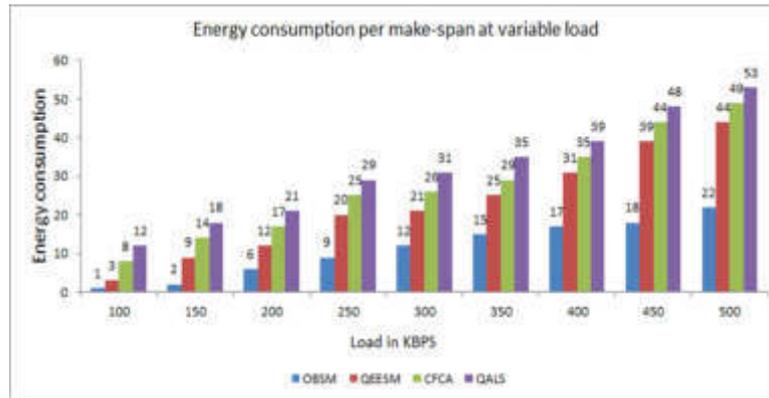


Fig. 4.4: The consumption of energy used for every make-span in variable load that was observed using the OBSM, QEESM, CFCA, & QALS methods

model for energy-efficient and sustainable Internet of things applications.

The reviewer recommended enhancing visual clarity, particularly in the flow chart (Figure 3.1), despite the study's strengths. To gain a deeper understanding of the system's functioning, a more comprehensive discussion of OBSM's algorithms and their impact on model performance is warranted.

**5. Conclusion.** The important problem of scheduling large-scale, time-sensitive data transmission over fog-enabled IoT networks in industry is addressed in this manuscript. Current scheduling models perform well for small to moderately delayed data transmissions, but not well for highly delayed data. This difference is closed by the fog-enabled IoT network-specific Optimal Batch Scheduling Model (OBSM). When it comes to performance, OBSM outperforms QEESM, CFCA, and QALS. OBSM's make-span rate is  $23.3 \pm 7.03$ , which shows that it is superior due to its lower average and minimal deviation. With  $17.8 \pm 5.2$ , OBSM beats QEESM ( $33.96 \pm 6.0$ ), CFCA ( $40.36 \pm 10.3$ ), and QALS ( $44.56 \pm 14.3$ ) in terms of average round-trip times. And in terms of energy efficiency, OBSM shines. Its energy consumption per make-span is lower ( $14.9 \pm 7.7$  joules) than that of QALS ( $31.8 \pm 12.8$ ), CFCA ( $27.4 \pm 12.8$ ), and QEESM ( $22.7 \pm 12.8$ ). These findings show how well the OBSM model performs in fog computing environments when handling large-scale, delay-sensitive data transmission. In terms of make-span rate, round-trip time, and energy consumption ratio, OBSM performs better than current models. Further studies will look more closely at IoT network task and workflow scheduling enabled by fog computing. Enhancing data transmission strategies in these crucial technological ecosystems requires this.

## REFERENCES

- [1] C. PANIAGUA AND J. DELSING, *Industrial frameworks for internet of things: A survey*, IEEE Systems Journal, 15 (2020), pp. 1149–1159.
- [2] T. QIU, J. CHI, X. ZHOU, Z. NING, M. ATIQUZZAMAN, AND D. O. WU, *Edge computing in industrial internet of things: Architecture, advances and challenges*, IEEE Communications Surveys & Tutorials, 22 (2020), pp. 2462–2488.
- [3] W. WU, Z. ZHAO, L. SHEN, X. T. R. KONG, D. GUO, R. Y. ZHONG, AND G. Q. HUANG, *Just Trolley: Implementation of industrial IoT and digital twin-enabled spatial-temporal traceability and visibility for finished goods logistics*, Advanced Engineering Informatics, 52 (2022), pp. 101571.
- [4] S. ALAM, S. T. SIDDIQUI, A. AHMAD, R. AHMAD, AND M. SHUAIB, *Internet of things (IoT) enabling technologies, requirements, and security challenges*, in Advances in Data and Information Sciences: Proceedings of ICDIS 2019, Springer Singapore, 2020, pp. 119–126.
- [5] F. AMIN, R. ABBASI, A. MATEEN, M. A. ABID, AND S. KHAN, *A step toward next-generation advancements in the internet of things technologies*, Sensors, 22 (2022), pp. 8072.
- [6] M. BANSAL, M. NANDA, AND M. N. HUSAIN, *Security and privacy aspects for Internet of Things (IoT)*, in 2021 6th International Conference on Inventive Computation Technologies (ICICT), IEEE, 2021, pp. 199–204.
- [7] B. C. CSÁJI, Z. KEMÉNY, G. PEDONE, A. KUTI, AND J. VÁNCZA, *Wireless multi-sensor networks for smart cities: A prototype system with statistical data analysis*, IEEE Sensors J., 17 (2017), pp. 7667–7676.

- [8] X. LIU, N. XIONG, W. LI, AND Y. XIE, *An optimization scheme of adaptive dynamic energy consumption based on joint network-channel coding in wireless sensor networks*, IEEE Sensors J., 15 (2015), pp. 5158–5168.
- [9] Y. DUAN, W. LI, X. FU, Y. LUO, AND L. YANG, *A methodology for reliability of WSN based on software defined network in adaptive industrial environment*, IEEE/CAA J. Automatica Sinica, 5 (2018), pp. 74–82.
- [10] C. ZHU, Z. SHENG, V. C. M. LEUNG, L. SHU, AND L. T. YANG, *Toward offering more useful data reliably to mobile cloud from wireless sensor network*, IEEE Trans. Emerg. Topics Comput., 3 (2015), pp. 84–94.
- [11] Y. WANG, W. YANG, X. SHANG, J. HU, Y. HUANG, AND Y. CAI, *Energy-efficient secure transmission for wireless powered Internet of things with multiple power beacons*, IEEE Access, 6 (2018), pp. 75086–75098.
- [12] S. KIM, C. KIM, AND J. KIM, *Reliable smart energy IoT-cloud service operation with container orchestration*, in Proc. 19th Asia-Pacific Netw. Oper. Manage. Symp. (APNOMS), Sep. 2017, pp. 378–381.
- [13] S. ZOPPI, A.V. BEMTEN, H. M. GÜRSU, M. VILGELM, J. GUCK, AND W. KELLERER, *Achieving hybrid wired/wireless industrial networks with WDetServ: Reliability-based scheduling for delay guarantees*, IEEE Trans. Ind. Inform., 14 (2018), no. 5, pp. 2307–2319.
- [14] Y. PENG, W. QIAO, L. QU, AND J. WANG, *Sensor fault detection and isolation for a wireless sensor network-based remote wind turbine condition monitoring system*, IEEE Trans. Ind. Appl., 54 (2018), no. 2, pp. 1072–1079.
- [15] W. SUN, W. LU, Q. LI, L. CHEN, D. MU, AND X. YUAN, *WNN-LQE: Wavelet neural-network-based link quality estimation for smart grid WSNs*, IEEE Access, 5 (2017), pp. 12788–12797.
- [16] J. YUAN AND X. LI, *A reliable and lightweight trust computing mechanism for IoT edge devices based on multi-source feedback information fusion*, IEEE Access, 6 (2018), pp. 23626–23638.
- [17] H. PARK, H. KIM, K. T. KIM, S.-T. KIM, AND P. MAH, *Frame-type-aware static time slotted channel hopping scheduling scheme for large-scale smart metering networks*, IEEE Access, 7 (2018), pp. 2200–2209.
- [18] S.-C. WANG, S.-C. TSENG, K.-Q. YAN, AND Y.-T. TSAI, *Reaching agreement in an integrated fog cloud IoT*, IEEE Access, 6 (2018), pp. 64515–64524.
- [19] D. PURKOVIC, M. HONSCH, AND T. R. M. K. MEYER, *An energy efficient communication protocol for low power, energy harvesting sensor modules*, IEEE Sensors J., 19 (2019), no. 2, pp. 701–714.
- [20] J. LUO ET AL., *Container-based fog computing architecture and energy-balancing scheduling algorithm for energy IoT*, Future Generation Computer Systems, 97 (2019), pp. 50–60.
- [21] H. Q. AL-SHAMMARI, A. LAWEY, T. EL-GORASHI, AND J. M. ELMIRGHANI, *Energy efficient service embedding in IoT networks*, in Proc. 27th Wireless Opt. Commun. Conf. (WOCC), Apr./May 2018, pp. 1–5.
- [22] I. S. M. ISA, M. O. I. MUSA, T. E. H. EL-GORASHI, A. Q. LAWEY, AND J. M. H. ELMIRGHANI, *Energy efficiency of fog computing health monitoring applications*, in Proc. 20th Int. Conf. Transparent Opt. Netw. (ICTON), Jul. 2018, pp. 1–5.
- [23] P. GAZORI, D. RAHBARI, M. NICKRAY, *Saving time and cost on the scheduling of fog-based IoT applications using deep reinforcement learning approach*, Future Generation Computer Systems, 110 (2020), pp. 1098–1115.
- [24] A. RAJITH, S. SOKI, AND M. HIROSHI, *Real-time optimized HVAC control system on top of an IoT framework*, in Proc. Third Int. Conf. Fog Mobile Edge Comput. (FMEEC), Apr. 2018, pp. 181–186.
- [25] A. LASZKA, S. EISELE, A. DUBEY, G. KARSAI, AND K. KVATERNIK, *TRANSAX: A blockchain-based decentralized forward-trading energy exchanged for transactive microgrids*, in Proc. IEEE 24th Int. Conf. Parallel Distrib. Syst. (ICPADS), Dec. 2018, pp. 918–927.
- [26] M. E. KHANOUCHE, Y. AMIRAT, A. CHIBANI, M. KERKAR, AND A. YACHIR, *Energy-centered and QoS-aware services selection for Internet of Things*, IEEE Trans. Autom. Sci. Eng., 13 (2016), no. 3, pp. 1256–1269.
- [27] K. RAMANA, ET AL., *Leaf disease classification in smart agriculture using deep neural network architecture and IoT*, Journal of Circuits, Systems and Computers, 31 (2022), no. 15, 2240004.
- [28] M. R. KUMAR, B. R. DEVI, K. RANGASWAMY, M. SANGEETHA AND K. V. R. KUMAR, *IoT-Edge Computing for Efficient and Effective Information Process on Industrial Automation*, 2023 International Conference on Networking and Communications (ICNWC), Chennai, India, 2023, pp. 1–6.
- [29] POTUNARAYANA, SREEDHARBHUKYA, CHANDRASHEKARJATOTH, P. PREMCHAND, *Quality-aware Energy Efficient Scheduling Model for Fog Enabled IoT Network*, Computers and Electrical Engineering, 2021.
- [30] V. K. A. KUMAR, M. R. KUMAR, N. SHRIBALA, N. SINGH, V. K. GUNJAN, K. N.-E.-A. SIDDIQUEE, AND M. ARIF, *Dynamic Wavelength Scheduling by Multiobjectives in OBS Networks*, Journal of Mathematics, 2022, Article ID 3806018.
- [31] J. LUO, ET AL., *Container-based fog computing architecture and energy-balancing scheduling algorithm for energy IoT*, Future Generation Computer Systems, 97 (2019), pp. 50–60.
- [32] S. M. VADLAMAANI, P. K. BHARTI, AND M. R. KUMAR, *Generalized Statistical Indicators For Cloud Computing Fault Tolerance*, International Journal of Intelligent Systems and Applications in Engineering, 10 (2022), no. 1s, pp. 269–281.
- [33] M. BHATIA, S. K. SOOD, AND S. KAUR, *Quantumized approach of load scheduling in fog computing environment for IoT applications*, Computing, (2020), pp. 1–19.
- [34] A. M. FAHIM, ET AL., *An efficient enhanced k-means clustering algorithm*, Journal of Zhejiang University-Science A, 7 (2006), no. 10, pp. 1626–1633.

*Edited by:* Anil Kumar Budati

*Special issue on:* Soft Computing and Artificial Intelligence for wire/wireless Human-Machine Interface

*Received:* Sep 24, 2023

*Accepted:* Dec 23, 2023



## TRAJECTORY INTERCEPTION CLASSIFICATION FOR PREDICTION OF COLLISION SCOPE BETWEEN MOVING OBJECTS

B. UMA MAHESH BABU\*, K. GIRI BABU† AND B. T. KRISHNA‡

**Abstract.** In the fields of autonomous navigation and vehicle safety, accurately predicting potential collision field points between moving objects is a significant challenge. A novel computing technique to enhance trajectory interception analysis is presented in this paper. Our objective is to develop a field model that can accurately forecast collision zones, improving road transportation safety and the use of autonomous cars. Our main contribution is a binary classification model called PCSMO (Prediction of Collision Scope between Moving Objects), which is based on zero-shot learning. Gann angles, which are typically 45 degrees, are used to analyze the trajectories of moving objects. This method is inspired by GANN (Gann Angle Numeric Nomenclature). Compared to earlier techniques, this model more accurately identifies potential collision interception zones. The technique computes Gann angles for trajectory analysis and extracts GPS coordinates of moving objects from video data using OpenCV. It offers a more sophisticated comprehension of object movement patterns and points of interception. To assess the precision, recall, F1-score, and prediction accuracy of our model, we employ 10-fold cross-validation. Comparing the PCSMO model to existing models, these metrics demonstrate how well the PCSMO model predicts potential collision zones. Our approach, we discovered, enhances trajectory analysis—a critical component of safer autonomous navigation systems. With potential applications in autonomous vehicle and UAV safety, the PCSMO model improves field interception classification.

**Key words:** Collision, Moving Objects. Global Positioning System, Machine Learning, Binary Classification, Gann Angle Degree, Trajectory Interception Detection, Unmanned Aerial Vehicles, Zero-Shot Learning.

**1. Introduction.** A conglomeration of advanced technological systems in combination with contemporary mobility solutions has created a paradigm shift in the ways supply chain ecosystem, and commutation systems across the world are becoming safer. There is a various set of tools, technical, management practices, and instrumentation engineering practices that are paving for safety in the mobility systems. Right from a two-wheeler to the jumbo-jet airliners at every level, the reliance on technology solutions has raised notches, and today, there are scores of control models that guard safe transportation and mobility.

Today, the contemporary practices of Unmanned Aerial Vehicles (UAVs), self-driven cars, drone-based delivery chains, robotic solutions, and AI solutions manning the traffic monitoring and control systems refer to a paradigm shift in futuristic solutions. While the scope of new-age solutions looks promising, still the scope for enhancements to the security and overall efficiency of autonomous vehicles is imperative [1].

Increasing demand for autonomous vehicles and vehicles with sensors for traffic mobility is on the rise, the systems must be more equipped in terms of mechanisms, patterns in which the systems are being deployed, and the measures that can help in improving the overall process of drive-safe conditions. In addition to the road-safety conditions with driver-less vehicles, even in the case the unmanned aerial vehicles, the role of systems in predicting the projectile path and the possible cross in the trajectory is impeccable need. As the domain is gaining traction, and the need for more comprehensive solutions are imperative, in this manuscript, the scope of developing prediction models that can be resourceful in the trajectory path crossing conditions is explored.

Projectile path trajectories estimation models are significant in the domain, and there are numerous models that were developed which could address the patterns and possible intersections. One key area wherein the studies are limited is the scope of motion analysis pertaining to a greater number of projectiles being in its motion simultaneously. However, there are a certain set of fundamental mathematical models and physics-theories-related equations and algorithms proposed earlier, which can address the projectile path conditions [2], [3]. Developing a secondary system that can create affirmative to the primary analysis approach can be a

---

\*ECE department, JNTUK Kakinada; Andhra Pradesh, India (Corresponding author, [umamahesh.ballla@gmail.com](mailto:umamahesh.ballla@gmail.com))

†Electronics and Communication Department in VVIT, Guntur, Andhra Pradesh ([giribabu@yahoo.com](mailto:giribabu@yahoo.com)).

‡ECE department, JNTUK Kakinada; Andhra Pradesh, India ([tkbattula@gmail.com](mailto:tkbattula@gmail.com)).

significant solution and embracing such solutions can lead to significant benefits for the predictive models, which can be resourceful in autonomous vehicles or UAV movement too. In this manuscript, focus is on the scope of continuous objective movement tracking and estimation of any close areas wherein the potential trajectory cross, and early stage of predictions of such trajectories [4]. Following are the key objectives considered in development of the proposed model:

- The first objective is to track the projectile path of an object in motion and its trajectory
- The second objective is to understand how the projectile path of an object can interfere with the other object in motion that is being tracked.
- Classification of the possible interceptions according to threat levels is based on the dynamic movement of the projectiles.

Targeting the above objectives, the approach in the model is to develop a systematic model wherein the key paths used by the projectiles are monitored and accordingly develop an interception mapping system that can change according to the real-time environment.

**2. Related Work.** Scores of models pertaining to object collisions on path prediction, accidents-related risk mitigation in autonomous vehicle segments, traffic density-based accident prediction zones, potential paths of gliding, and intersections are some of the actual ranges of studies explored in the literature review. From the summation of key points in the literature, the following are some of the critical observations and learnings: The majority of the trajectory-related studies or safety-related studies are reliant on third-party components or tools like sensors, GPS positioning tools, density meters, altitude meters, speedometers, etc. The performance of the whole system is dependent on the inputs attained from the equipment or tools adapted in the models. Irrespective of the efficacy of the models, any depletion in the quality of the tools used for gauging metrics could be a serious threat in terms of miscalculation and ineffective predictions of the intersections [5–11].

The other key observation is about application of the solutions for univariate analysis, like one moving object movement compared to the other static object sensors, etc. When there are more than one object making moves in zigzag or varying directions, the complexities of accuracy are high for the predictive models [12–16].

“COLLIDE-PRED,” developed by Author et al. [17], uses the motion of objects moving to provide collision predictions. It is a pipeline that begins with object identification, which is utilized for object tracking; afterwards, trajectory prediction is conducted, which culminates in collision detection. The authors indicated that the COLLIDE-PRED will determine the probable site of the collision. This model attempted to determine the target object’s trajectory’s collision scope, which often exhibits false alarm. In addition, the collision scope of moving objects can only be predicted using this approach for offline video streams.

Applying the machine learning model is seen as a promising solution from the literature towards testing the label classification approach and other such metrics, which are significant for the execution of the models. There are scales of statistical and technology-centric models discussed in the literature towards understanding the scope of autonomous vehicle safety in its trajectory. However, there is a need for more explorative studies in the domain to improve the overall efficiency with which the solutions can be managed. In order to address the constraints addressed in this review of contemporary literature, this manuscript suggests a novel zero-shot binary classification model that discovers threat zones and safe zones of trajectory interceptions.

In the further sections of this manuscript, Section 2 refers to the related work summary from the literature review. Section 3 provides insights into the materials and methods, the proposed model narrative, its algorithm flow, and other key metrics that signify the mode. Section 4 provides insights into the experimental study, and Section 5 refers to the conclusion based on the efficiency aspects estimated from the model.

**3. Proposed Method and Materials.** The method of classifying trajectory interceptions with and without collision scope between moving objects and respective materials have been explored in this section. The section includes the narrative of the suggested model. By uniquely combining zero-shot learning with GANN principles, the PCSMO (Prediction of Collision Scope between Moving Objects) predictive model presents a novel approach to collision detection. Unlike traditional models, the PCSMO model can reliably predict collision points without historical data thanks to this integration. The key to this innovation is the application of Gann angles, which are used in financial markets, to trajectory analysis. These angles provide a new perspective on trajectory interception classification by enabling the PCSMO model to predict collision points in dynamic

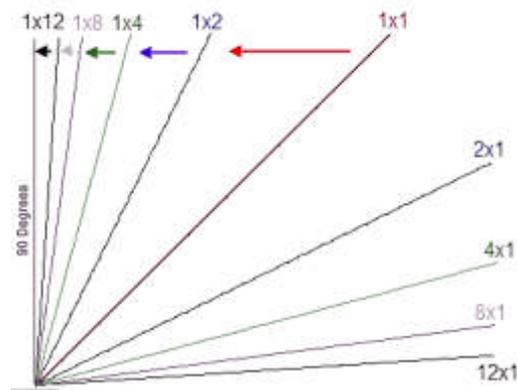


Fig. 3.1: The Gann angle degree

scenarios such as road traffic and autonomous vehicle navigation. In order to improve Gann angle calculations, the model additionally makes use of OpenCV to extract precise GPS coordinates from video data. This approach adds something uniquely to the field of trajectory analysis while improving the predictive power of the model.

**3.1. Model Narrative.** The model proposed in this manuscript is based on the trajectory detection requirements, and fundamentally the concept of Gann Theory is used in the development of the model. While the metrics and techniques used in the assessment of the path and motion are prevalent in many of the earlier studies, application of GANN angle formations is a unique concept explored in this manuscript.

Gann Angles are named after its creator W.D. Gann, and the theory is used across the verticals for various set of analysis. Profoundly known for its application in the securities and commodities price feasibility analysis, the intensity, and the logic behind the application of Gann theories across various industrial principles and practices are phenomenal. Gann angle approach is based on x-degree angle detection from the origin or base point used for estimation, wherein important trends are detected.

Applying the concept of Gann, the model proposed in this manuscript is to use the object-based sensors to detect the distance from the primary object to the tracking object, and using the latitude and longitude details of the tracking object (to), the Gann angle degree shall be developed Figure 3.1.

Based on the Gann angle detected, each angle border is coded in a sequential manner, and the motion of the projectile or the object is tracked between one specific angle ranges. Thus, for every “n” period in further the motion of the object is tracked for identifying the possible angle in which the trajectory could be forming. In similar manner, even for the primary object (Po) too, the motion is captured based on the latitude and longitude and the possible projectile of the object is developed [18].

The interception method adapted here is to find the degree of path for both the objects (Po and To) and identify any interception areas between the Gann angles. For instance, if there are three major areas wherein the object crossing each other is identified, it shall be narrowed down as a ‘potential interception area’ in motion.

The process is repeated for every “n” period, depending on the speed, height or other relative metrics as adapted in the earlier studies. However, such metrics shall be used for deciding the appropriate time frame for measuring the motion analysis. Based on the number of interception layers in the angles forming as diamonds or squares, such areas can be seen as sensitive zone for interception.

Depending on the motion of both the primary and secondary objects, the projectile path interceptions could alter based on the fresh Gann Angles. In the initial Gann Angles, if new interception location is formed, and no over-lapping is formed, such zones shall be removed from the Threat Zone count, and the new ones are added. If there is any over-lapping of the zones to the earlier identified zones, then the classification of possible threat of interception in such a zone shall be upgraded.

By adapting such a system, the scope of analysis in terms of path can be easier and it leads to minimalistic

approach in the detection process. For instance, if the speed or the height or traffic or other metrics are used as the primary parameters, there could be huge modifications to such metrics in a real-time environment. Whereas, when the process is based on an angle detection, despite of small fluctuations in the speed or obstructions in the path etc., still the path and feasibility of the movement is broad-lined, and at every “n” period when the analysis is carried out, it refers to the multiple inputs like variance in angles from the origin, and the possible areas wherein the interception can take place.

If the purpose is to find the exact location and time of the trajectory cross is the objective, there are many existing complex solutions. But to identify the potential zone of threat, relying on the Gann Angle approach can lead to sustainable outcome.

**3.2. Rationale of the Approach.** More often the trajectory assessments are carried out for two different aspects. One to understand the intersection points, and the secondary towards predicting the trajectory path, wherein necessary action to thwart any such challenges can take place. However, considering certain metrics like momentum, obstructions, speed, weight, path glide, etc. multiple sets of elements are to be tracked in terms of predicting the trajectory and interceptions, which could increase the complexities of computation [7].

Whereas, in the preliminary level, a simple angle-based assessment of the movement refers to the possible line of movements for the object. For instance, when an object is driving in a direction, the angles at x-degrees are drawn to either direction of movement. Thus, there is a clear path projection as to between what angles the object is heading on. In the following “n” period, when the reassessment of the latitude and longitude is taking place, it refers to the angular shift or continuity taking place. Accordingly, the angles can be extended to the “direction of the movement”.

Similarly, when both the objects like the primary object and the tracking objects movement angles are decided, it shall help in mitigating the risks and improving the potential interception areas classification. Adapting to such patterns, can help in reducing the load on different set of metrics being captured, and thus mitigating the complexities pertaining to the process.

**3.3. Primary and Target Object.** The object in the context of the proposed model could be something which is in movement. Irrespective of the direction and the path, speed or other metrics, the object could be presumed for an autonomous vehicle or UAV or even a human driven vehicle etc. The classification in terms of primary and target objects is as follows.

The primary object is the object which is in the control of the monitoring team and being controlled for its maneuvering in a specific direction. On contrary, the target object is the secondary object which is being targeted for understanding the path and projectile. For tracking the target object and the primary target, the need for understanding the latitude and longitude position is paramount. The whole process of analysis shall be based on the latitude and longitude assessment which can be procured based on the GPS trackers [3], [8].

**3.4. GPS Trackers.** A global positioning system, as a tracker, is a device installed on an object, and any positional changes to the object are tracked over a map. Based on real-time tracking, the tracking of factors like latitude and longitude, the direction of movements can easily be tracked for an application system [6].

**3.5. Latitude and Longitude.** Latitude and longitude coordinates are intended for determining and describing the position as well as location of any point on the Earth’s surface.

**3.6. Interception Points..** In the motion of an object, there is a certain path in which the object moves depending on the speed, path etc. In the trajectory planned for the process, there could be certain areas wherein two or more objects in its path could be colliding and such points of collision can be seen as interception points. In the context of the proposed model, the interception points shall be formed based on the overlapping Gann points in the angles derived from the directions in which the Gann angles are formed for both the Primary Object and the Target Object [9], [12].

**3.7. “n” Period.** “n” be the notional period represented here wherein n can be a specific period for which the next course of latitude and longitude analysis and the direction of the objects of interest in motion is targeted. Unless the “n” period is computed, identifying the next overlapping intercept points is not feasible. Thus, there is a need for appropriate levels of estimating the n-period angle of movements.

**3.8. Threat Zone Count.** The threat zone count is the summation of a number of squares or block overlaps intercepted with two overlapping zones of Gann Angles related to the primary and target objectives. Depending on the number of additions and eliminations of the overlaps, for each period “n” a new number of threat zone counts shall be imperative for one primary object in motion to the target objects.

In furtherance, the same concept can be applied to multiple object movements too to trace the potential areas of trajectory interceptions, and that can help in addressing the possible movements and path [13].

**3.9. The classification strategy.** The suggested model performs zero-shot learning [19] based on binary classification of trajectory interceptions of the moving objects such that, they are prone to collision or not. Unlike conventional machine learning-based classification strategies, the zero-shot learning approach performs classification without a training phase. However, the zero-shot learning model can perform classification on both seen class as well as unseen class data [20]. The suggested model performs binary classification of the fewer seen class data. Here, the fewer seen class data [21] denotes the combination of seen and unseen class data. The objective function derived to perform the classification of trajectory interceptions is capable to classify the trajectory zones as a safe zone, threat zone found, threat zone with less possibility of collision, threat zone with moderate possibility of collision, and threat zone with the certainty of collision. As a result, based on the classification result, the machine intelligence can alert the driving forces of the moving objects such that:

If the class predicted is a safe zone then no alert to driving forces.

If the class predicted is a threat zone, an alert with no action recommendation to driving forces.

If the class predicted is a threat zone with less collision possibility (yellow zone), then results in a continuous alert about collision scope to driving forces.

If the class predicted is a threat zone with moderate collision possibility (orange zone), then results in a continuous alert of a recommended action (such as slow down the object, or trajectory track change) about collision scope to driving forces.

If the class predicted is a threat zone with high collision possibility (red zone), then results from a continuous alert about collision scope with recommended action (stop the object, if not slow down and change the trajectory tracking of the object).

**3.10. The Data.** The data used in experiments have been synthesized from the compilations of the trajectory interceptions of the moving vehicles captured on cc cameras, which have publicly been available on YouTube [22]. Overall, 1000 images were captured from these compilations. Among these, 215 captures were pruned due to a lack of sensitivity and specificity in trajectory interception observed. The rest of the captures were annotated as having trajectory interceptions prone to collision (positive class), and safe trajectory interceptions (negative class). The count of positive class and negative class captures are 391 and 394 respectively. These resultant captures were preprocessed using OpenCV [23] resulting in GPS coordinates of the sources of moving vehicles. The further phase discovers the further GPS coordinates of the trajectories of the corresponding vehicles. Afterward, determines the Gann-inspired angles for each of the GPS coordinates of the corresponding vehicles. For each capture of both positive and negative classes, a set of records will be framed in the format projected in the following figure (figure 3.2). The projected figure (figure 3.2) indicating the record format of the processed dataset representing the both classes positive and negative such that, for each object  $V_i$ , represents set of GPS coordinates  $[C_1, C_2, \dots, C_m]$ , and for each coordinate  $C_j$  discovers set of Gann angles based zones  $[GZ_1, GZ_2, \dots, GZ_p]$

**3.11. The Objective Function of Collision Prediction.** This section explores the objective function of the proposed binary classification strategy to predict trajectory interceptions with collision scope between two moving objects.

**3.11.1. Model Definition..** In order to predict the scope of trajectory interception between the given two moving objects considered, the key phases involved are projected in the following description. For a given two objects stated as primary object  $P$  and target object  $T$

Initial step of the suggested model discovers all GPS coordinates  $P(L), T(L)$  of the moving directions of the respective objects. Following, for each GPS coordinate of the both objects, discovers GANN angles. Later, the suggested model discovers trajectories of the both objects. Further for each trajectory  $P_t$  of the primary object  $P$ , verifies the scope of interception between the trajectory  $P_t$  of primary and each trajectory  $T_t$  of target

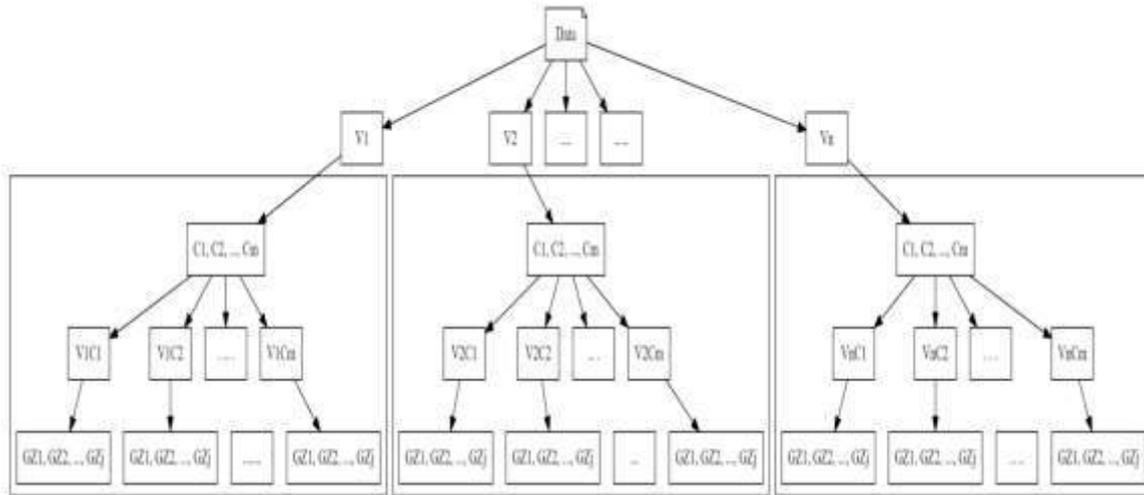


Fig. 3.2: Each capture of both positive and negative classes

object as follows If GPS coordinates of the both objects under the present trajectory of primary and target objects same, then determines, Gann angle based zones indexed in clockwise direction for both primary and secondary objects. If the trajectory of primary object and trajectory of target object are sharing the same index of Gann angle zone, then the respective zone stated as threat zone. If the criteria meets at subsequent GPS coordinate, then threat zone states as threat zone-yellow If criteria meets at further GPS coordinates, then the corresponding threat zone will be stated as threat zone-orange and threat zone-red in respective order.

**3.12. Mathematical Modeling of the Approach.**

**1. Definitions:**

- $P$  // primary object
- $T$  // target object
- $P(L)$  // series of GPS coordinates observed for primary object  $P$
- $T(L)$  // series of GPS coordinates observed for target object  $T$ , which is as follows:
- $D$  // total duration of the objects in motion
- $u$  // the unit of time

**2. GPS Coordinate Tracking:**

- $i = 1$
- while  $(i \cdot u) \leq D$  //Begin
  - $P(L) \leftarrow \text{GPS}(P)$  // The Eq 1 present GPS coordinates of primary object
  - $T(L) \leftarrow \text{GPS}(T)$  // The Eq 2 present GPS coordinates of target object
  - $i \leftarrow i + 1$
- End

**3. Gann Angle Series Creation:**

- For each  $g_i \in P(L)$ 
  - $d^0 = 0^\circ$  // present angle of the Gann series
  - $S^\circ(g_i)$  // series of angles, which is an empty set
  - while  $d^\circ < 360^\circ$  Begin
    - \*  $d^\circ \leftarrow d^\circ + x^\circ$
    - \*  $S^\circ(g_i) \leftarrow d^\circ$
  - End

**4. Collision Scope Detection:**

- Let  $aI$  be the angle index that meets the criteria  $1 \leq aI \leq |S^\circ(g_i)|$  and representing GANN angles

in clockwise direction. Begin: // Eq 3

- For each  $g_i \in P(L) \cap T(L)$ 
  - \* The object's trajectory (Zone Angle of Moving Object) of primary and target objects denoted as primary trajectory  $P_t$  and target trajectory  $T_t$  referred by the expressions  $P_t^{aI}$  and  $T_t^{aI}$ .
  - \* if  $(aI \notin \text{thrZ} \wedge (P_t^{aI} \cap T_t^{aI}) \neq \emptyset)$  Begin
    - $\text{thrZ} \leftarrow aI$  // Eq 4
  - \* End
- Repeat for other conditions and equations as required.

- End: collisionScopeDetection

#### 5. Additional Steps and Details:

- The stated 'collisionScopeDetection' repeats if there is a change in angle Index  $aI$ .
- Additional steps and details as required.

**4. Experimental Study Analysis.** This section details the experimental study carried out to assess the performance of the suggested zero-shot learning based binary classification of trajectory interceptions as prone to collision or safe. The 10-fold cross validation was adopted to scale the precision, specificity, sensitivity and accuracy of the suggested classification model. In order to exhibit the performance advantage of the suggested model, the values exhibited for cross validation metrics were compared to the cross validation metric values exhibited by the contemporary model COLLIDE-PRED [17]. The comparative study of the resultant cross validation metric values of the proposed and contemporary models exhibiting that the proposed model PCSMO is outperforming the contemporary model towards prediction of collision scope between moving objects. Python was used to implement the proposed approach [24], and the code was built using the Python editor PyCharm [25]. In this regard, I5-7th gen Intel processor with 32 GB of memory and a 1TB storage was considered for the hardware requirements.

Gann Theory is applied in the PCSMO trajectory detection model. The foundation of the novelty of our model is the Gann Angle, named after W.D. Gann, which is widely applied in commodities and securities price feasibility analysis. These methods are specifically applied to trajectory interception classification in our model, a novel application that has not been explored before. The x-degree angles are the particular parameters that are obtained from Gann Angles and are crucial in identifying movement trends starting from an origin or base point. The application of these angles enables our model to predict interception points by identifying overlapping Gann angles, which are crucial for determining the potential collision zones between the Primary Object and the Target Object. It is necessary to compute intercept points using the 'n' period parameter, which establishes a time frame for trajectory analysis. Without this temporal component, potential overlapping intercept points between moving objects cannot be identified. The significance of this parameter for the temporal analysis of object trajectories led to its selection. Threat Zone Count is another important parameter in the PCSMO algorithm. A quantitative indicator of the potential collision scope is the sum of squares or block overlaps within two overlapping Gann Angle zones. Since this parameter physically represents collision risk areas, we use it to evaluate the PCSMO algorithm in our experimental study. Finally, we conclude that the theoretical significance of our experimental parameters for Gann Theory, their analytical power in other domains, and their empirical significance in trajectory analysis and collision prediction within the PCSMO model, guided their selection. The statistics of the fewer seen class data that detailed in section 3.10 are exhibited in following table (table 4.1).

The ratio of true positive predictions to made positive predictions is known as precision, or positive predictive value. This metric is critical when the cost of false positives is high. In terms of accurate collision scenario prediction without overprediction, the PCSMO model performed better than COLLIDE-PRED.

The precision of the cross-validation measure relates to the positive predictive value acquired by cross-validation from both the proposed PCSMO and existing COLLIDE-PRED methods, respectively. Figure 4.1 depicts the precision data obtained from each cross-validation fold. The positive predictive value (PPV) acquired by the proposed PCSMO method is larger than that obtained by COLLIDE-PRED, implying that the proposed method can effectively reduce false negative rates. In comparison to the current model, COLLIDE-PRED, the metric precision values obtained show the relevance of the PCSMO with low variance. PCSMO shows a higher

Table 4.1: The statistics of the fewer seen class data

Parameter	Count
No of captures from source compilations	1000
No of captures considered for fewer seen class data preparation	785
No of captures listed as positive class (trajectory interceptions prone to collision)	391
No of captures listed as negative class (safe trajectory interceptions)	394
No of positive class records	3152
No of negative class records	3519
No of unseen records	672

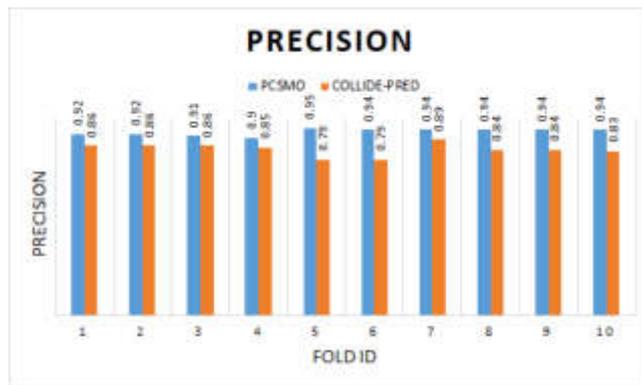


Fig. 4.1: The positive predictive values observed from cross-validation

mean precision (0.930) compared to COLLIDE-PRED (0.841). This indicates that PCSMO is more reliable in its positive predictions, with fewer false positives on average.

The percentage of real negatives that are correctly identified is known as the True Negative Rate, or Specificity. This displays the accuracy of the model’s non-collision detection. Indicating PCSMO’s dependability in locating safe trajectories, it showed high specificity and low variance.

The True Negative Rate, which corresponds to the specificity of the PCSMO and COLLIDE-PRED techniques, is the result of tenfold cross-validation. Figure 4.2 summarizes the specificity acquired from each of the 10-folds of the cross-validation. The numbers obtained for metric specificity indicate the relevance of the PCSMO with minimum variation from the current model, COLLIDE-PRED. PCSMO has a higher mean specificity (0.927) than COLLIDE-PRED (0.837), suggesting that PCSMO is better at correctly identifying negative cases without falsely categorizing them as positive.

They are balanced by the F-Measure, which is the harmonic mean of recall (sensitivity) and precision. Particularly helpful in cases of unequal class distribution. Indicating a balanced performance between sensitivity and precision, the f-measure of the PCSMO model was significantly higher than COLLIDE-PRED.

For the recommended and contemporaneous models, the cross-validation metric "F-measure" was determined among 10 folds, as shown in Figure 4.3. In descending order, the average f-measures for PCSMO and COLLIDE-PRED are displayed. The comprehensive statistics of the F-measure generated from tenfold cross-validation are shown in Figure 4.3. The metric f-measure values obtained demonstrate that the PCSMO is more relevant than the current model COLLIDE-PRED in terms of performance. The proposed model PCSMO outperforms the present model COLLIDE-PRED, as evidenced by the substantial F-measure value produced from the recommended model PCSMO. The mean F-measure for PCSMO (0.927) exceeds COLLIDE-PRED’s (0.839), which suggests a better balance between precision and sensitivity in PCSMO, making it a more robust model overall. The ability of the model to recognize every positive instance is known as its sensitivity, or true



Fig. 4.2: The True Negative Rate obtained from cross-validation

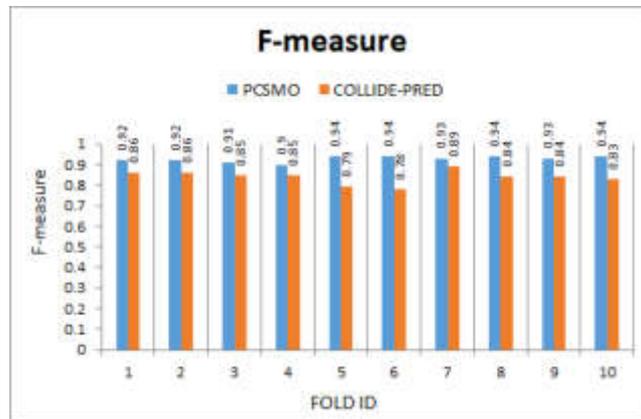


Fig. 4.3: The f-measure observations from tenfold cross-validation

positive rate. The PCSMO model demonstrated exceptional sensitivity in locating possible collision points across the majority of folds, despite having a lower true positive rate.

The true positive rate also refers to the sensitivity discovered by cross-validation on both PCSMO and COLLIDE-PRED techniques, respectively. Figure 4.4 depicts the fold level sensitivity in further detail. When comparing the PCSMO to the current model COLLIDE-PRED, the metric sensitivity values show that the PCSMO outperforms the COLLIDE-PRED. Despite having a lower true positive rate, the suggested PCSMO model appears to have superior sensitivity in the majority of folds. PCSMO's mean sensitivity (0.949) is greater than that of COLLIDE-PRED (0.874). PCSMO is more effective at correctly identifying true positive cases, indicating a lower miss rate for actual positives. The percentage of records correctly classified is called accuracy. It is a crucial performance metric for models in classification problems. The PCSMO model predicted collision and non-collision scenarios more accurately than the COLLIDE-PRED model in the majority of cross-validation folds.

The "accuracy" is a cross-validation statistic that measures the proportion of properly categorized records to the total number of records. PCSMO and COLLIDE-PRED both have good overall prediction accuracy. Figure 4.5 shows the accuracy statistics derived from tenfold cross-validation in more detail. When comparing the PCSMO to the current model COLLIDE-PRED, the metric accuracy results show that the PCSMO outperforms the COLLIDE-PRED. The accuracy distribution chart for cross-validation of the PCSMO as well as COLLIDE-PRED models is shown in Figure 4.5. The mean accuracy of PCSMO (0.939) is significantly higher than that

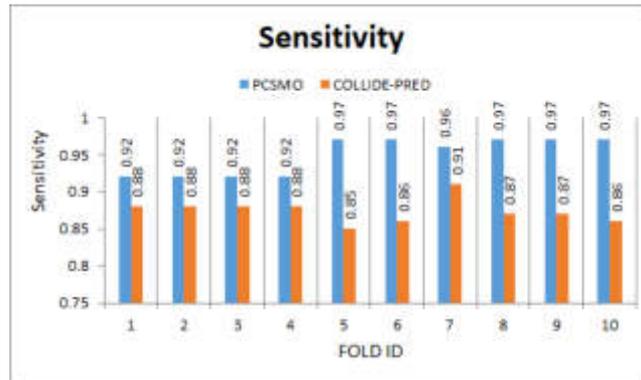


Fig. 4.4: The sensitivity observed from tenfold cross-validation

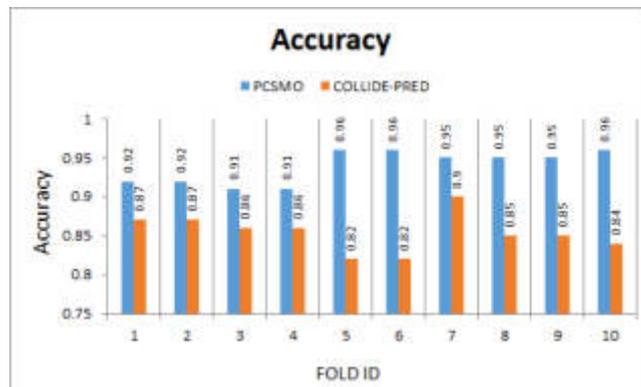


Fig. 4.5: Accuracy obtained from tenfold cross-validation

of COLLIDE-PRED (0.854), meaning that PCSMO correctly predicts both positive and negative outcomes more often than COLLIDE-PRED. In binary classification problems, the Matthews Correlation Coefficient (MCC) provides more information than accuracy when classes have significant differences in size. It evaluates negatives and true versus false positives and is thought to be balanced. PCSMO's MCC values exceeded COLLIDE-PRED's, demonstrating the predictive ability of the model.

The cross-validation metric MCC has been measured over the proposed and contemporary models among ten folds, as shown in Figure 4.6. The MCC observed for PCSMO and COLLIDE-PRED in the respective order. The detailed statistics of MCC obtained from tenfold cross-validation have been explored in Figure 4.6. The values obtained for metric MCC exhibits the significance of the PCSMO with better performance while compared to the contemporary model COLLIDE-PRED. PCSMO's mean MCC (0.875) is notably higher than COLLIDE-PRED's (0.714). Since MCC is a balanced measure even when the classes are of different sizes, a higher MCC indicates a better performance of PCSMO, with a stronger correlation between observed and predicted classifications.

**5. Conclusion.** In conclusion, it can be challenging to evaluate trajectories and locate intersections where moving objects might collide, particularly in road transportation where cars can travel in several directions. Potential collision zones between moving objects are predicted by Zero-shot Learning-based Trajectory Interception Classification. The potential collision locations and parameters between the primary and target objects are determined using angle-based analysis, which is incorporated into our PCSMO model. Threat zones are defined by overlapping Gann angles. After a thorough 10-fold cross-validation, the model's efficacy was determined.

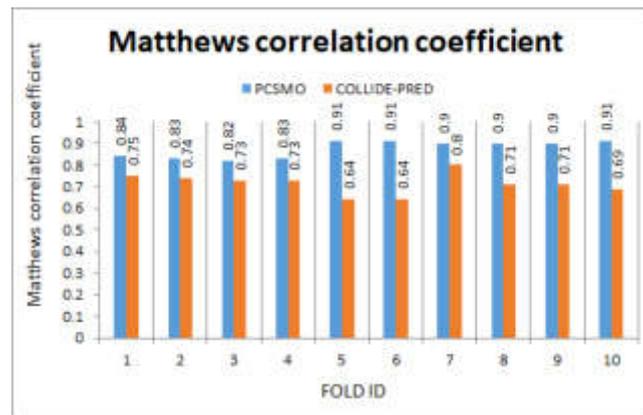


Fig. 4.6: The Mathews Correlation Coefficient (MCC) observed from the tenfold cross-validation

It fared significantly better than the COLLIDE-PRED model in every performance metric, with a prediction accuracy above 90%. From wide to narrow, the PCSMO model identifies the angle range and width of potential collision zones. The micro-level collision point is not precisely located by it. The Gann Nine scale approach will be added to this model in order to improve collision scope predictions. By using this refined scale, we hope to more accurately define the threat range and provide a second-level analysis that could significantly lower the risk of collisions. By defining threat zones, the Gann Nine square model extension will help us focus our predictive skills.

#### REFERENCES

- [1] J. CHEN, ET AL., *Path Planning for Autonomous Vehicle Based on a Two-Layered Planning Model in Complex Environment*, Journal of Advanced Transportation, 2020.
- [2] S. ZHU AND B. AKSUN-GUVENC, *Trajectory planning of autonomous vehicles based on parameterized control optimization in dynamic on-road environments*, Journal of Intelligent & Robotic Systems, vol. 100, no. 3, pp. 1055-1067, 2020.
- [3] C. KATRAKAZAS, ET AL., *Real-time motion planning methods for autonomous on-road driving: State-of-the-art and future research directions*, Transportation Research Part C: Emerging Technologies, vol. 60, pp. 416-442, 2015.
- [4] J. NILSSON, ET AL., *Manoeuvre generation and control for automated highway driving*, IFAC Proceedings Volumes, vol. 47, no. 3, pp. 6301-6306, 2014.
- [5] J. CHOU AND J. J. TASY, *Trajectory planning of autonomous vehicles in environments with moving obstacles*, IFAC Proceedings Volumes, vol. 26, no. 1, pp. 439-445, 1993.
- [6] F. ZHANG, R. XIA, AND X. CHEN, *An optimal trajectory planning algorithm for autonomous trucks: architecture, algorithm, and experiment*, International Journal of Advanced Robotic Systems, vol. 17, no. 2, 1729881420909603, 2020.
- [7] D. YANG, ET AL., *A dynamic lane-changing trajectory planning model for automated vehicles*, Transportation Research Part C: Emerging Technologies, vol. 95, pp. 228-247, 2018.
- [8] L. LI, J. LI, AND S. ZHANG, *State-of-the-art trajectory tracking of autonomous vehicles*, Mechanical Sciences, vol. 12, no. 1, pp. 419-432, 2021.
- [9] A. EINARSSON, *Composing music as an embodied activity*, Editors: A. Arweström Jansson, A. Axelsson, R. Andreasson, Uppsala University, 2017.
- [10] S. BORIC, ET AL., *Research in Autonomous Driving—A Historic Bibliometric View of the Research Development in Autonomous Driving*, International Journal of Innovation and Economic Development, vol. 7, no. 5, pp. 27-44, 2021.
- [11] C. WANG, ET AL., *Trajectory planning for autonomous vehicles based on improved Hybrid A*, International Journal of Vehicle Design, vol. 83, no. 2-4, pp. 218-239, 2020.
- [12] S. MANOHARAN, *An improved safety algorithm for artificial intelligence enabled processors in self driving cars*, Journal of artificial intelligence, vol. 1, no. 02, pp. 95-104, 2019.
- [13] A. GOSWAMI, *Trajectory generation for lane-change maneuver of autonomous vehicles*, Dissertation, Purdue University, 2015.
- [14] H. KHAYYAM, ET AL., *Artificial intelligence and internet of things for autonomous vehicles*, in Nonlinear approaches in engineering applications, Springer, Cham, pp. 39-68, 2020.
- [15] I. A. KULIKOV, I. A. ULCHENKO, AND A. V. CHAPLYGIN, *Using real world data in virtual development and testing of a path tracking controller for an autonomous vehicle*, International Journal of Innovative Technology and Exploring Engineering (IJITEE), vol. 8, no. 12, pp. 720-726, 2019.

- [16] A. ARTUNEDO, J. VILLAGRA, AND J. GODOY, *Real-time motion planning approach for automated driving in urban environments*, IEEE Access, vol. 7, pp. 180039-180053, 2019.
- [17] D. CHAVAN, D. SAAD, AND D. B. CHAKRABORTY, *COLLIDE-PRED: Prediction of On-Road Collision From Surveillance Videos*, arXiv preprint arXiv:2101.08463, 2021.
- [18] NIRMAL BANG, *What is Gann Theory?*, available at <https://www.nirmalbang.com/knowledge-center/what-is-gann-theory.html>.
- [19] S. CHANGPINYO, ET AL., *Synthesized classifiers for zero-shot learning*, Proceedings of the IEEE conference on computer vision and pattern recognition, 2016.
- [20] R. SOCHER, ET AL., *Zero-shot learning through cross-modal transfer*, Advances in neural information processing systems 26, 2013.
- [21] V. K. VERMA, ET AL., *Towards zero-shot learning with fewer seen class examples*, Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2021.
- [22] *Missing vehicle collisions recorded on CCTV compilation*, available at [https://www.youtube.com/results?search\\_query=missing+vehicle+collisions+recorded+on+cctv+compilation](https://www.youtube.com/results?search_query=missing+vehicle+collisions+recorded+on+cctv+compilation).
- [23] G. BRADSKI AND A. KAEHLER, *OpenCV*, Dr. Dobb's journal of software tools, vol. 3, 2000.
- [24] *Python*, available at <https://www.python.org/downloads/>.
- [25] *Pycharm*, available at <https://www.jetbrains.com/pycharm/download/>.

*Edited by:* Anil Kumar Budati

*Special issue on:* Soft Computing and Artificial Intelligence for wire/wireless Human-Machine Interface

*Received:* Sep 24, 2023

*Accepted:* Dec 23, 2023



## TRAJECTORY INTERCEPTION CLASSIFICATION FOR PREDICTION OF COLLISION SCOPE BETWEEN MOVING OBJECTS

B. UMA MAHESH BABU\*, K. GIRI BABU† AND B. T. KRISHNA‡

**Abstract.** In the fields of autonomous navigation and vehicle safety, accurately predicting potential collision field points between moving objects is a significant challenge. A novel computing technique to enhance trajectory interception analysis is presented in this paper. Our objective is to develop a field model that can accurately forecast collision zones, improving road transportation safety and the use of autonomous cars. Our main contribution is a binary classification model called PCSMO (Prediction of Collision Scope between Moving Objects), which is based on zero-shot learning. Gann angles, which are typically 45 degrees, are used to analyze the trajectories of moving objects. This method is inspired by GANN (Gann Angle Numeric Nomenclature). Compared to earlier techniques, this model more accurately identifies potential collision interception zones. The technique computes Gann angles for trajectory analysis and extracts GPS coordinates of moving objects from video data using OpenCV. It offers a more sophisticated comprehension of object movement patterns and points of interception. To assess the precision, recall, F1-score, and prediction accuracy of our model, we employ 10-fold cross-validation. Comparing the PCSMO model to existing models, these metrics demonstrate how well the PCSMO model predicts potential collision zones. Our approach, we discovered, enhances trajectory analysis—a critical component of safer autonomous navigation systems. With potential applications in autonomous vehicle and UAV safety, the PCSMO model improves field interception classification.

**Key words:** Collision, Moving Objects. Global Positioning System, Machine Learning, Binary Classification, Gann Angle Degree, Trajectory Interception Detection, Unmanned Aerial Vehicles, Zero-Shot Learning.

**1. Introduction.** A conglomeration of advanced technological systems in combination with contemporary mobility solutions has created a paradigm shift in the ways supply chain ecosystem, and commutation systems across the world are becoming safer. There is a various set of tools, technical, management practices, and instrumentation engineering practices that are paving for safety in the mobility systems. Right from a two-wheeler to the jumbo-jet airliners at every level, the reliance on technology solutions has raised notches, and today, there are scores of control models that guard safe transportation and mobility.

Today, the contemporary practices of Unmanned Aerial Vehicles (UAVs), self-driven cars, drone-based delivery chains, robotic solutions, and AI solutions manning the traffic monitoring and control systems refer to a paradigm shift in futuristic solutions. While the scope of new-age solutions looks promising, still the scope for enhancements to the security and overall efficiency of autonomous vehicles is imperative [1].

Increasing demand for autonomous vehicles and vehicles with sensors for traffic mobility is on the rise, the systems must be more equipped in terms of mechanisms, patterns in which the systems are being deployed, and the measures that can help in improving the overall process of drive-safe conditions. In addition to the road-safety conditions with driver-less vehicles, even in the case the unmanned aerial vehicles, the role of systems in predicting the projectile path and the possible cross in the trajectory is impeccable need. As the domain is gaining traction, and the need for more comprehensive solutions are imperative, in this manuscript, the scope of developing prediction models that can be resourceful in the trajectory path crossing conditions is explored.

Projectile path trajectories estimation models are significant in the domain, and there are numerous models that were developed which could address the patterns and possible intersections. One key area wherein the studies are limited is the scope of motion analysis pertaining to a greater number of projectiles being in its motion simultaneously. However, there are a certain set of fundamental mathematical models and physics-theories-related equations and algorithms proposed earlier, which can address the projectile path conditions

---

\*ECE department, JNTUK Kakinada; Andhra Pradesh, India (Corresponding author, [umamahesh.ballala@gmail.com](mailto:umamahesh.ballala@gmail.com))

†Electronics and Communication Department in VVIT, Guntur, Andhra Pradesh ([giribabu@yahoo.com](mailto:giribabu@yahoo.com)).

‡ECE department, JNTUK Kakinada; Andhra Pradesh, India ([tkbattula@gmail.com](mailto:tkbattula@gmail.com)).

[2], [3]. Developing a secondary system that can create affirmative to the primary analysis approach can be a significant solution and embracing such solutions can lead to significant benefits for the predictive models, which can be resourceful in autonomous vehicles or UAV movement too. In this manuscript, focus is on the scope of continuous objective movement tracking and estimation of any close areas wherein the potential trajectory cross, and early stage of predictions of such trajectories [4]. Following are the key objectives considered in development of the proposed model:

- The first objective is to track the projectile path of an object in motion and its trajectory
- The second objective is to understand how the projectile path of an object can interfere with the other object in motion that is being tracked.
- Classification of the possible interceptions according to threat levels is based on the dynamic movement of the projectiles.

Targeting the above objectives, the approach in the model is to develop a systematic model wherein the key paths used by the projectiles are monitored and accordingly develop an interception mapping system that can change according to the real-time environment.

**2. Related Work.** Scores of models pertaining to object collisions on path prediction, accidents-related risk mitigation in autonomous vehicle segments, traffic density-based accident prediction zones, potential paths of gliding, and intersections are some of the actual ranges of studies explored in the literature review. From the summation of key points in the literature, the following are some of the critical observations and learnings: The majority of the trajectory-related studies or safety-related studies are reliant on third-party components or tools like sensors, GPS positioning tools, density meters, altitude meters, speedometers, etc. The performance of the whole system is dependent on the inputs attained from the equipment or tools adapted in the models. Irrespective of the efficacy of the models, any depletion in the quality of the tools used for gauging metrics could be a serious threat in terms of miscalculation and ineffective predictions of the intersections [5–11].

The other key observation is about application of the solutions for univariate analysis, like one moving object movement compared to the other static object sensors, etc. When there are more than one object making moves in zigzag or varying directions, the complexities of accuracy are high for the predictive models [12–16].

“COLLIDE-PRED,” developed by Author et al. [17], uses the motion of objects moving to provide collision predictions. It is a pipeline that begins with object identification, which is utilized for object tracking; afterwards, trajectory prediction is conducted, which culminates in collision detection. The authors indicated that the COLLIDE-PRED will determine the probable site of the collision. This model attempted to determine the target object’s trajectory’s collision scope, which often exhibits false alarm. In addition, the collision scope of moving objects can only be predicted using this approach for offline video streams.

Applying the machine learning model is seen as a promising solution from the literature towards testing the label classification approach and other such metrics, which are significant for the execution of the models. There are scales of statistical and technology-centric models discussed in the literature towards understanding the scope of autonomous vehicle safety in its trajectory. However, there is a need for more explorative studies in the domain to improve the overall efficiency with which the solutions can be managed. In order to address the constraints addressed in this review of contemporary literature, this manuscript suggests a novel zero-shot binary classification model that discovers threat zones and safe zones of trajectory interceptions.

In the further sections of this manuscript, Section 2 refers to the related work summary from the literature review. Section 3 provides insights into the materials and methods, the proposed model narrative, its algorithm flow, and other key metrics that signify the mode. Section 4 provides insights into the experimental study, and Section 5 refers to the conclusion based on the efficiency aspects estimated from the model.

**3. Proposed Method and Materials.** The method of classifying trajectory interceptions with and without collision scope between moving objects and respective materials have been explored in this section. The section includes the narrative of the suggested model. By uniquely combining zero-shot learning with GANN principles, the PCSMO (Prediction of Collision Scope between Moving Objects) predictive model presents a novel approach to collision detection. Unlike traditional models, the PCSMO model can reliably predict collision points without historical data thanks to this integration. The key to this innovation is the application of Gann angles, which are used in financial markets, to trajectory analysis. These angles provide a new perspective

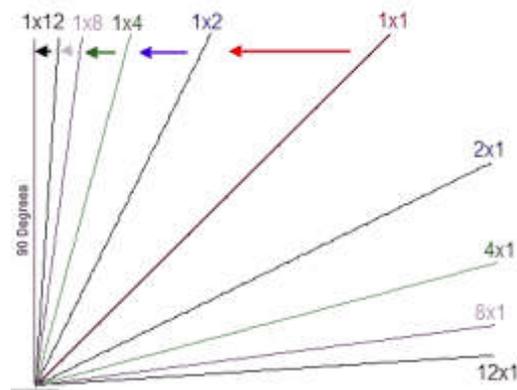


Fig. 3.1: The Gann angle degree

on trajectory interception classification by enabling the PCSMO model to predict collision points in dynamic scenarios such as road traffic and autonomous vehicle navigation. In order to improve Gann angle calculations, the model additionally makes use of OpenCV to extract precise GPS coordinates from video data. This approach adds something uniquely to the field of trajectory analysis while improving the predictive power of the model.

**3.1. Model Narrative.** The model proposed in this manuscript is based on the trajectory detection requirements, and fundamentally the concept of Gann Theory is used in the development of the model. While the metrics and techniques used in the assessment of the path and motion are prevalent in many of the earlier studies, application of GANN angle formations is a unique concept explored in this manuscript.

Gann Angles are named after its creator W.D. Gann, and the theory is used across the verticals for various set of analysis. Profoundly known for its application in the securities and commodities price feasibility analysis, the intensity, and the logic behind the application of Gann theories across various industrial principles and practices are phenomenal. Gann angle approach is based on x-degree angle detection from the origin or base point used for estimation, wherein important trends are detected.

Applying the concept of Gann, the model proposed in this manuscript is to use the object-based sensors to detect the distance from the primary object to the tracking object, and using the latitude and longitude details of the tracking object (to), the Gann angle degree shall be developed Figure 3.1.

Based on the Gann angle detected, each angle border is coded in a sequential manner, and the motion of the projectile or the object is tracked between one specific angle ranges. Thus, for every “n” period in further the motion of the object is tracked for identifying the possible angle in which the trajectory could be forming. In similar manner, even for the primary object (Po) too, the motion is captured based on the latitude and longitude and the possible projectile of the object is developed [18].

The interception method adapted here is to find the degree of path for both the objects (Po and To) and identify any interception areas between the Gann angles. For instance, if there are three major areas wherein the object crossing each other is identified, it shall be narrowed down as a ‘potential interception area’ in motion.

The process is repeated for every “n” period, depending on the speed, height or other relative metrics as adapted in the earlier studies. However, such metrics shall be used for deciding the appropriate time frame for measuring the motion analysis. Based on the number of interception layers in the angles forming as diamonds or squares, such areas can be seen as sensitive zone for interception.

Depending on the motion of both the primary and secondary objects, the projectile path interceptions could alter based on the fresh Gann Angles. In the initial Gann Angles, if new interception location is formed, and no over-lapping is formed, such zones shall be removed from the Threat Zone count, and the new ones are added. If there is any over-lapping of the zones to the earlier identified zones, then the classification of possible threat of interception in such a zone shall be upgraded.

By adapting such a system, the scope of analysis in terms of path can be easier and it leads to minimalistic approach in the detection process. For instance, if the speed or the height or traffic or other metrics are used as the primary parameters, there could be huge modifications to such metrics in a real-time environment. Whereas, when the process is based on an angle detection, despite of small fluctuations in the speed or obstructions in the path etc., still the path and feasibility of the movement is broad-lined, and at every “n” period when the analysis is carried out, it refers to the multiple inputs like variance in angles from the origin, and the possible areas wherein the interception can take place.

If the purpose is to find the exact location and time of the trajectory cross is the objective, there are many existing complex solutions. But to identify the potential zone of threat, relying on the Gann Angle approach can lead to sustainable outcome.

**3.2. Rationale of the Approach.** More often the trajectory assessments are carried out for two different aspects. One to understand the intersection points, and the secondary towards predicting the trajectory path, wherein necessary action to thwart any such challenges can take place. However, considering certain metrics like momentum, obstructions, speed, weight, path glide, etc. multiple sets of elements are to be tracked in terms of predicting the trajectory and interceptions, which could increase the complexities of computation [7].

Whereas, in the preliminary level, a simple angle-based assessment of the movement refers to the possible line of movements for the object. For instance, when an object is driving in a direction, the angles at x-degrees are drawn to either direction of movement. Thus, there is a clear path projection as to between what angles the object is heading on. In the following “n” period, when the reassessment of the latitude and longitude is taking place, it refers to the angular shift or continuity taking place. Accordingly, the angles can be extended to the “direction of the movement”.

Similarly, when both the objects like the primary object and the tracking objects movement angles are decided, it shall help in mitigating the risks and improving the potential interception areas classification. Adapting to such patterns, can help in reducing the load on different set of metrics being captured, and thus mitigating the complexities pertaining to the process.

**3.3. Primary and Target Object.** The object in the context of the proposed model could be something which is in movement. Irrespective of the direction and the path, speed or other metrics, the object could be presumed for an autonomous vehicle or UAV or even a human driven vehicle etc. The classification in terms of primary and target objects is as follows.

The primary object is the object which is in the control of the monitoring team and being controlled for its maneuvering in a specific direction. On contrary, the target object is the secondary object which is being targeted for understanding the path and projectile. For tracking the target object and the primary target, the need for understanding the latitude and longitude position is paramount. The whole process of analysis shall be based on the latitude and longitude assessment which can be procured based on the GPS trackers [3], [8].

**3.4. GPS Trackers.** A global positioning system, as a tracker, is a device installed on an object, and any positional changes to the object are tracked over a map. Based on real-time tracking, the tracking of factors like latitude and longitude, the direction of movements can easily be tracked for an application system [6].

**3.5. Latitude and Longitude.** Latitude and longitude coordinates are intended for determining and describing the position as well as location of any point on the Earth’s surface.

**3.6. Interception Points..** In the motion of an object, there is a certain path in which the object moves depending on the speed, path etc. In the trajectory planned for the process, there could be certain areas wherein two or more objects in its path could be colliding and such points of collision can be seen as interception points. In the context of the proposed model, the interception points shall be formed based on the overlapping Gann points in the angles derived from the directions in which the Gann angles are formed for both the Primary Object and the Target Object [9], [12].

**3.7. “n” Period.** “n” be the notional period represented here wherein n can be a specific period for which the next course of latitude and longitude analysis and the direction of the objects of interest in motion is targeted. Unless the “n” period is computed, identifying the next overlapping intercept points is not feasible. Thus, there is a need for appropriate levels of estimating the n-period angle of movements.

**3.8. Threat Zone Count.** The threat zone count is the summation of a number of squares or block overlaps intercepted with two overlapping zones of Gann Angles related to the primary and target objectives. Depending on the number of additions and eliminations of the overlaps, for each period “n” a new number of threat zone counts shall be imperative for one primary object in motion to the target objects.

In furtherance, the same concept can be applied to multiple object movements too to trace the potential areas of trajectory interceptions, and that can help in addressing the possible movements and path [13].

**3.9. The classification strategy.** The suggested model performs zero-shot learning [19] based on binary classification of trajectory interceptions of the moving objects such that, they are prone to collision or not. Unlike conventional machine learning-based classification strategies, the zero-shot learning approach performs classification without a training phase. However, the zero-shot learning model can perform classification on both seen class as well as unseen class data [20]. The suggested model performs binary classification of the fewer seen class data. Here, the fewer seen class data [21] denotes the combination of seen and unseen class data. The objective function derived to perform the classification of trajectory interceptions is capable to classify the trajectory zones as a safe zone, threat zone found, threat zone with less possibility of collision, threat zone with moderate possibility of collision, and threat zone with the certainty of collision. As a result, based on the classification result, the machine intelligence can alert the driving forces of the moving objects such that:

If the class predicted is a safe zone then no alert to driving forces.

If the class predicted is a threat zone, an alert with no action recommendation to driving forces.

If the class predicted is a threat zone with less collision possibility (yellow zone), then results in a continuous alert about collision scope to driving forces.

If the class predicted is a threat zone with moderate collision possibility (orange zone), then results in a continuous alert of a recommended action (such as slow down the object, or trajectory track change) about collision scope to driving forces.

If the class predicted is a threat zone with high collision possibility (red zone), then results from a continuous alert about collision scope with recommended action (stop the object, if not slow down and change the trajectory tracking of the object).

**3.10. The Data.** The data used in experiments have been synthesized from the compilations of the trajectory interceptions of the moving vehicles captured on cc cameras, which have publicly been available on YouTube [22]. Overall, 1000 images were captured from these compilations. Among these, 215 captures were pruned due to a lack of sensitivity and specificity in trajectory interception observed. The rest of the captures were annotated as having trajectory interceptions prone to collision (positive class), and safe trajectory interceptions (negative class). The count of positive class and negative class captures are 391 and 394 respectively. These resultant captures were preprocessed using OpenCV [23] resulting in GPS coordinates of the sources of moving vehicles. The further phase discovers the further GPS coordinates of the trajectories of the corresponding vehicles. Afterward, determines the Gann-inspired angles for each of the GPS coordinates of the corresponding vehicles. For each capture of both positive and negative classes, a set of records will be framed in the format projected in the following figure (figure 3.2). The projected figure (figure 3.2) indicating the record format of the processed dataset representing the both classes positive and negative such that, for each object  $V_i$ , represents set of GPS coordinates  $[C_1, C_2, \dots, C_m]$ , and for each coordinate  $C_j$  discovers set of Gann angles based zones  $[GZ_1, GZ_2, \dots, GZ_p]$

**3.11. The Objective Function of Collision Prediction.** This section explores the objective function of the proposed binary classification strategy to predict trajectory interceptions with collision scope between two moving objects.

**3.11.1. Model Definition..** In order to predict the scope of trajectory interception between the given two moving objects considered, the key phases involved are projected in the following description. For a given two objects stated as primary object  $P$  and target object  $T$

Initial step of the suggested model discovers all GPS coordinates  $P(L), T(L)$  of the moving directions of the respective objects. Following, for each GPS coordinate of the both objects, discovers GANN angles. Later, the suggested model discovers trajectories of the both objects. Further for each trajectory  $P_t$  of the primary object  $P$ , verifies the scope of interception between the trajectory  $P_t$  of primary and each trajectory  $T_t$  of target

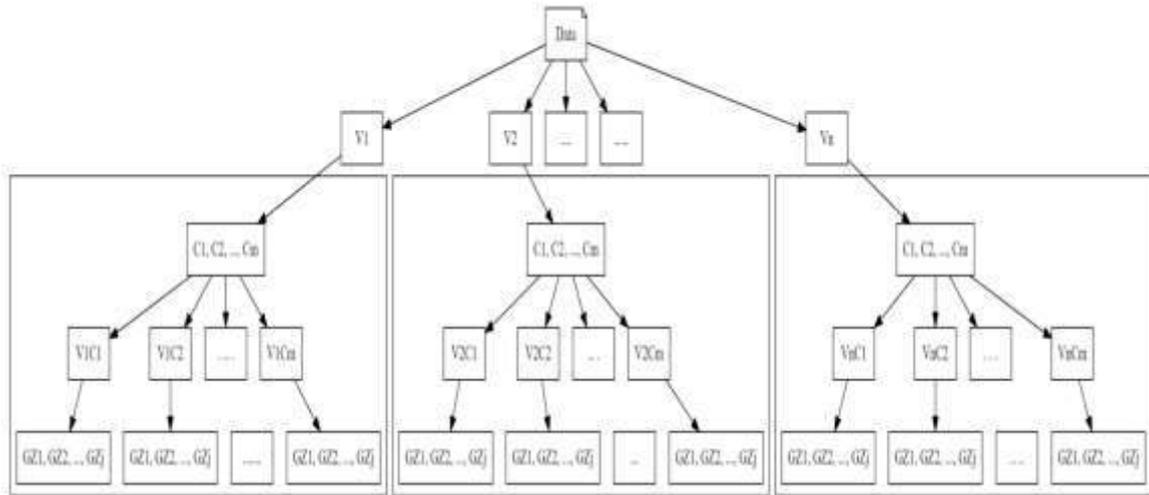


Fig. 3.2: Each capture of both positive and negative classes

object as follows If GPS coordinates of the both objects under the present trajectory of primary and target objects same, then determines, Gann angle based zones indexed in clockwise direction for both primary and secondary objects. If the trajectory of primary object and trajectory of target object are sharing the same index of Gann angle zone, then the respective zone stated as threat zone. If the criteria meets at subsequent GPS coordinate, then threat zone states as threat zone-yellow If criteria meets at further GPS coordinates, then the corresponding threat zone will be stated as threat zone-orange and threat zone-red in respective order.

**3.12. Mathematical Modeling of the Approach.**

**1. Definitions:**

- $P$  // primary object
- $T$  // target object
- $P(L)$  // series of GPS coordinates observed for primary object  $P$
- $T(L)$  // series of GPS coordinates observed for target object  $T$ , which is as follows:
- $D$  // total duration of the objects in motion
- $u$  // the unit of time

**2. GPS Coordinate Tracking:**

- $i = 1$
- while  $(i \cdot u) \leq D$  //Begin
  - $P(L) \leftarrow \text{GPS}(P)$  // The Eq 1 present GPS coordinates of primary object
  - $T(L) \leftarrow \text{GPS}(T)$  // The Eq 2 present GPS coordinates of target object
  - $i \leftarrow i + 1$
- End

**3. Gann Angle Series Creation:**

- For each  $g_i \in P(L)$ 
  - $d^0 = 0^\circ$  // present angle of the Gann series
  - $S^\circ(g_i)$  // series of angles, which is an empty set
  - while  $d^\circ < 360^\circ$  Begin
    - \*  $d^\circ \leftarrow d^\circ + x^\circ$
    - \*  $S^\circ(g_i) \leftarrow d^\circ$
  - End

**4. Collision Scope Detection:**

- Let  $aI$  be the angle index that meets the criteria  $1 \leq aI \leq |S^\circ(g_i)|$  and representing GANN angles

in clockwise direction. Begin: // Eq 3

- For each  $g_i \in P(L) \cap T(L)$ 
  - \* The object's trajectory (Zone Angle of Moving Object) of primary and target objects denoted as primary trajectory  $P_t$  and target trajectory  $T_t$  referred by the expressions  $P_t^{aI}$  and  $T_t^{aI}$ .
  - \* if  $(aI \notin \text{thrZ} \wedge (P_t^{aI} \cap T_t^{aI}) \neq \emptyset)$  Begin
    - $\text{thrZ} \leftarrow aI$  // Eq 4
  - \* End
- Repeat for other conditions and equations as required.

- End: collisionScopeDetection

#### 5. Additional Steps and Details:

- The stated 'collisionScopeDetection' repeats if there is a change in angle Index  $aI$ .
- Additional steps and details as required.

**4. Experimental Study Analysis.** This section details the experimental study carried out to assess the performance of the suggested zero-shot learning based binary classification of trajectory interceptions as prone to collision or safe. The 10-fold cross validation was adopted to scale the precision, specificity, sensitivity and accuracy of the suggested classification model. In order to exhibit the performance advantage of the suggested model, the values exhibited for cross validation metrics were compared to the cross validation metric values exhibited by the contemporary model COLLIDE-PRED [17]. The comparative study of the resultant cross validation metric values of the proposed and contemporary models exhibiting that the proposed model PCSMO is outperforming the contemporary model towards prediction of collision scope between moving objects. Python was used to implement the proposed approach [24], and the code was built using the Python editor PyCharm [25]. In this regard, I5-7th gen Intel processor with 32 GB of memory and a 1TB storage was considered for the hardware requirements.

Gann Theory is applied in the PCSMO trajectory detection model. The foundation of the novelty of our model is the Gann Angle, named after W.D. Gann, which is widely applied in commodities and securities price feasibility analysis. These methods are specifically applied to trajectory interception classification in our model, a novel application that has not been explored before. The x-degree angles are the particular parameters that are obtained from Gann Angles and are crucial in identifying movement trends starting from an origin or base point. The application of these angles enables our model to predict interception points by identifying overlapping Gann angles, which are crucial for determining the potential collision zones between the Primary Object and the Target Object. It is necessary to compute intercept points using the 'n' period parameter, which establishes a time frame for trajectory analysis. Without this temporal component, potential overlapping intercept points between moving objects cannot be identified. The significance of this parameter for the temporal analysis of object trajectories led to its selection. Threat Zone Count is another important parameter in the PCSMO algorithm. A quantitative indicator of the potential collision scope is the sum of squares or block overlaps within two overlapping Gann Angle zones. Since this parameter physically represents collision risk areas, we use it to evaluate the PCSMO algorithm in our experimental study. Finally, we conclude that the theoretical significance of our experimental parameters for Gann Theory, their analytical power in other domains, and their empirical significance in trajectory analysis and collision prediction within the PCSMO model, guided their selection. The statistics of the fewer seen class data that detailed in section 3.10 are exhibited in following table (table 4.1).

The ratio of true positive predictions to made positive predictions is known as precision, or positive predictive value. This metric is critical when the cost of false positives is high. In terms of accurate collision scenario prediction without overprediction, the PCSMO model performed better than COLLIDE-PRED.

The precision of the cross-validation measure relates to the positive predictive value acquired by cross-validation from both the proposed PCSMO and existing COLLIDE-PRED methods, respectively. Figure 4.1 depicts the precision data obtained from each cross-validation fold. The positive predictive value (PPV) acquired by the proposed PCSMO method is larger than that obtained by COLLIDE-PRED, implying that the proposed method can effectively reduce false negative rates. In comparison to the current model, COLLIDE-PRED, the metric precision values obtained show the relevance of the PCSMO with low variance. PCSMO shows a higher

Table 4.1: The statistics of the fewer seen class data

Parameter	Count
No of captures from source compilations	1000
No of captures considered for fewer seen class data preparation	785
No of captures listed as positive class (trajectory interceptions prone to collision)	391
No of captures listed as negative class (safe trajectory interceptions)	394
No of positive class records	3152
No of negative class records	3519
No of unseen records	672

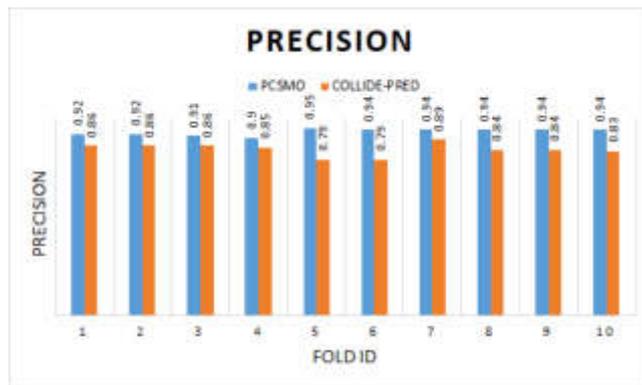


Fig. 4.1: The positive predictive values observed from cross-validation

mean precision (0.930) compared to COLLIDE-PRED (0.841). This indicates that PCSMO is more reliable in its positive predictions, with fewer false positives on average.

The percentage of real negatives that are correctly identified is known as the True Negative Rate, or Specificity. This displays the accuracy of the model’s non-collision detection. Indicating PCSMO’s dependability in locating safe trajectories, it showed high specificity and low variance.

The True Negative Rate, which corresponds to the specificity of the PCSMO and COLLIDE-PRED techniques, is the result of tenfold cross-validation. Figure 4.2 summarizes the specificity acquired from each of the 10-folds of the cross-validation. The numbers obtained for metric specificity indicate the relevance of the PCSMO with minimum variation from the current model, COLLIDE-PRED. PCSMO has a higher mean specificity (0.927) than COLLIDE-PRED (0.837), suggesting that PCSMO is better at correctly identifying negative cases without falsely categorizing them as positive.

They are balanced by the F-Measure, which is the harmonic mean of recall (sensitivity) and precision. Particularly helpful in cases of unequal class distribution. Indicating a balanced performance between sensitivity and precision, the f-measure of the PCSMO model was significantly higher than COLLIDE-PRED.

For the recommended and contemporaneous models, the cross-validation metric "F-measure" was determined among 10 folds, as shown in Figure 4.3. In descending order, the average f-measures for PCSMO and COLLIDE-PRED are displayed. The comprehensive statistics of the F-measure generated from tenfold cross-validation are shown in Figure 4.3. The metric f-measure values obtained demonstrate that the PCSMO is more relevant than the current model COLLIDE-PRED in terms of performance. The proposed model PCSMO outperforms the present model COLLIDE-PRED, as evidenced by the substantial F-measure value produced from the recommended model PCSMO. The mean F-measure for PCSMO (0.927) exceeds COLLIDE-PRED’s (0.839), which suggests a better balance between precision and sensitivity in PCSMO, making it a more robust model overall. The ability of the model to recognize every positive instance is known as its sensitivity, or true



Fig. 4.2: The True Negative Rate obtained from cross-validation

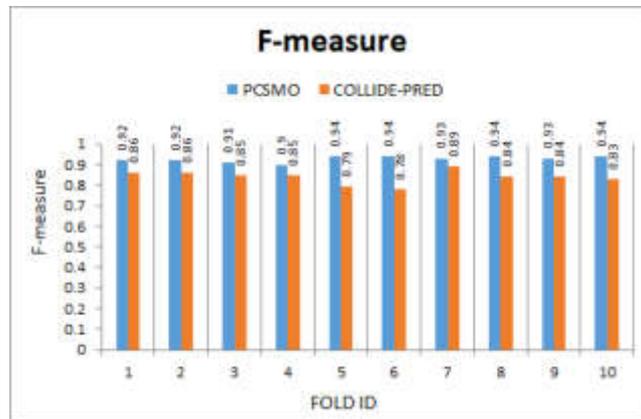


Fig. 4.3: The f-measure observations from tenfold cross-validation

positive rate. The PCSMO model demonstrated exceptional sensitivity in locating possible collision points across the majority of folds, despite having a lower true positive rate.

The true positive rate also refers to the sensitivity discovered by cross-validation on both PCSMO and COLLIDE-PRED techniques, respectively. Figure 4.4 depicts the fold level sensitivity in further detail. When comparing the PCSMO to the current model COLLIDE-PRED, the metric sensitivity values show that the PCSMO outperforms the COLLIDE-PRED. Despite having a lower true positive rate, the suggested PCSMO model appears to have superior sensitivity in the majority of folds. PCSMO's mean sensitivity (0.949) is greater than that of COLLIDE-PRED (0.874). PCSMO is more effective at correctly identifying true positive cases, indicating a lower miss rate for actual positives. The percentage of records correctly classified is called accuracy. It is a crucial performance metric for models in classification problems. The PCSMO model predicted collision and non-collision scenarios more accurately than the COLLIDE-PRED model in the majority of cross-validation folds.

The "accuracy" is a cross-validation statistic that measures the proportion of properly categorized records to the total number of records. PCSMO and COLLIDE-PRED both have good overall prediction accuracy. Figure 4.5 shows the accuracy statistics derived from tenfold cross-validation in more detail. When comparing the PCSMO to the current model COLLIDE-PRED, the metric accuracy results show that the PCSMO outperforms the COLLIDE-PRED. The accuracy distribution chart for cross-validation of the PCSMO as well as COLLIDE-PRED models is shown in Figure 4.5. The mean accuracy of PCSMO (0.939) is significantly higher than that

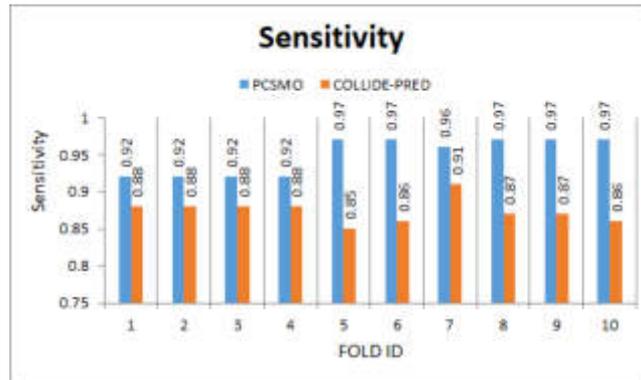


Fig. 4.4: The sensitivity observed from tenfold cross-validation

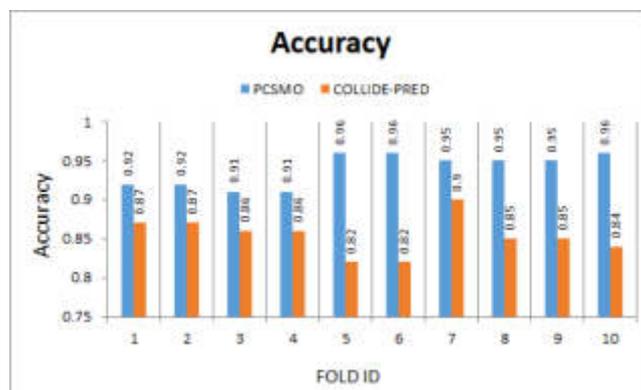


Fig. 4.5: Accuracy obtained from tenfold cross-validation

of COLLIDE-PRED (0.854), meaning that PCSMO correctly predicts both positive and negative outcomes more often than COLLIDE-PRED. In binary classification problems, the Matthews Correlation Coefficient (MCC) provides more information than accuracy when classes have significant differences in size. It evaluates negatives and true versus false positives and is thought to be balanced. PCSMO's MCC values exceeded COLLIDE-PRED's, demonstrating the predictive ability of the model.

The cross-validation metric MCC has been measured over the proposed and contemporary models among ten folds, as shown in Figure 4.6. The MCC observed for PCSMO and COLLIDE-PRED in the respective order. The detailed statistics of MCC obtained from tenfold cross-validation have been explored in Figure 4.6. The values obtained for metric MCC exhibits the significance of the PCSMO with better performance while compared to the contemporary model COLLIDE-PRED. PCSMO's mean MCC (0.875) is notably higher than COLLIDE-PRED's (0.714). Since MCC is a balanced measure even when the classes are of different sizes, a higher MCC indicates a better performance of PCSMO, with a stronger correlation between observed and predicted classifications.

**5. Conclusion.** In conclusion, it can be challenging to evaluate trajectories and locate intersections where moving objects might collide, particularly in road transportation where cars can travel in several directions. Potential collision zones between moving objects are predicted by Zero-shot Learning-based Trajectory Interception Classification. The potential collision locations and parameters between the primary and target objects are determined using angle-based analysis, which is incorporated into our PCSMO model. Threat zones are defined by overlapping Gann angles. After a thorough 10-fold cross-validation, the model's efficacy was determined.

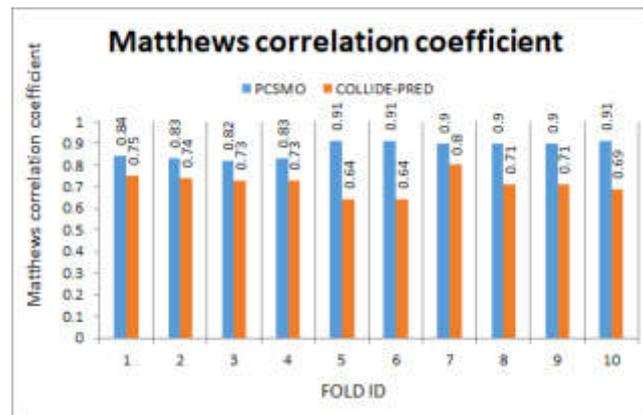


Fig. 4.6: The Mathews Correlation Coefficient (MCC) observed from the tenfold cross-validation

It fared significantly better than the COLLIDE-PRED model in every performance metric, with a prediction accuracy above 90%. From wide to narrow, the PCSMO model identifies the angle range and width of potential collision zones. The micro-level collision point is not precisely located by it. The Gann Nine scale approach will be added to this model in order to improve collision scope predictions. By using this refined scale, we hope to more accurately define the threat range and provide a second-level analysis that could significantly lower the risk of collisions. By defining threat zones, the Gann Nine square model extension will help us focus our predictive skills.

#### REFERENCES

- [1] J. CHEN, ET AL., *Path Planning for Autonomous Vehicle Based on a Two-Layered Planning Model in Complex Environment*, Journal of Advanced Transportation, 2020.
- [2] S. ZHU AND B. AKSUN-GUVENC, *Trajectory planning of autonomous vehicles based on parameterized control optimization in dynamic on-road environments*, Journal of Intelligent & Robotic Systems, vol. 100, no. 3, pp. 1055-1067, 2020.
- [3] C. KATRAKAZAS, ET AL., *Real-time motion planning methods for autonomous on-road driving: State-of-the-art and future research directions*, Transportation Research Part C: Emerging Technologies, vol. 60, pp. 416-442, 2015.
- [4] J. NILSSON, ET AL., *Manoeuvre generation and control for automated highway driving*, IFAC Proceedings Volumes, vol. 47, no. 3, pp. 6301-6306, 2014.
- [5] J. CHOU AND J. J. TASY, *Trajectory planning of autonomous vehicles in environments with moving obstacles*, IFAC Proceedings Volumes, vol. 26, no. 1, pp. 439-445, 1993.
- [6] F. ZHANG, R. XIA, AND X. CHEN, *An optimal trajectory planning algorithm for autonomous trucks: architecture, algorithm, and experiment*, International Journal of Advanced Robotic Systems, vol. 17, no. 2, 1729881420909603, 2020.
- [7] D. YANG, ET AL., *A dynamic lane-changing trajectory planning model for automated vehicles*, Transportation Research Part C: Emerging Technologies, vol. 95, pp. 228-247, 2018.
- [8] L. LI, J. LI, AND S. ZHANG, *State-of-the-art trajectory tracking of autonomous vehicles*, Mechanical Sciences, vol. 12, no. 1, pp. 419-432, 2021.
- [9] A. EINARSSON, *Composing music as an embodied activity*, Editors: A. Arweström Jansson, A. Axelsson, R. Andreasson, Uppsala University, 2017.
- [10] S. BORIC, ET AL., *Research in Autonomous Driving—A Historic Bibliometric View of the Research Development in Autonomous Driving*, International Journal of Innovation and Economic Development, vol. 7, no. 5, pp. 27-44, 2021.
- [11] C. WANG, ET AL., *Trajectory planning for autonomous vehicles based on improved Hybrid A*, International Journal of Vehicle Design, vol. 83, no. 2-4, pp. 218-239, 2020.
- [12] S. MANOHARAN, *An improved safety algorithm for artificial intelligence enabled processors in self driving cars*, Journal of artificial intelligence, vol. 1, no. 02, pp. 95-104, 2019.
- [13] A. GOSWAMI, *Trajectory generation for lane-change maneuver of autonomous vehicles*, Dissertation, Purdue University, 2015.
- [14] H. KHAYYAM, ET AL., *Artificial intelligence and internet of things for autonomous vehicles*, in Nonlinear approaches in engineering applications, Springer, Cham, pp. 39-68, 2020.
- [15] I. A. KULIKOV, I. A. ULCHENKO, AND A. V. CHAPLYGIN, *Using real world data in virtual development and testing of a path tracking controller for an autonomous vehicle*, International Journal of Innovative Technology and Exploring Engineering (IJITEE), vol. 8, no. 12, pp. 720-726, 2019.

- [16] A. ARTUNEDO, J. VILLAGRA, AND J. GODOY, *Real-time motion planning approach for automated driving in urban environments*, IEEE Access, vol. 7, pp. 180039-180053, 2019.
- [17] D. CHAVAN, D. SAAD, AND D. B. CHAKRABORTY, *COLLIDE-PRED: Prediction of On-Road Collision From Surveillance Videos*, arXiv preprint arXiv:2101.08463, 2021.
- [18] NIRMAL BANG, *What is Gann Theory?*, available at <https://www.nirmalbang.com/knowledge-center/what-is-gann-theory.html>.
- [19] S. CHANGPINYO, ET AL., *Synthesized classifiers for zero-shot learning*, Proceedings of the IEEE conference on computer vision and pattern recognition, 2016.
- [20] R. SOCHER, ET AL., *Zero-shot learning through cross-modal transfer*, Advances in neural information processing systems 26, 2013.
- [21] V. K. VERMA, ET AL., *Towards zero-shot learning with fewer seen class examples*, Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2021.
- [22] *Missing vehicle collisions recorded on CCTV compilation*, available at [https://www.youtube.com/results?search\\_query=missing+vehicle+collisions+recorded+on+cctv+compilation](https://www.youtube.com/results?search_query=missing+vehicle+collisions+recorded+on+cctv+compilation).
- [23] G. BRADSKI AND A. KAEHLER, *OpenCV*, Dr. Dobb's journal of software tools, vol. 3, 2000.
- [24] *Python*, available at <https://www.python.org/downloads/>.
- [25] *Pycharm*, available at <https://www.jetbrains.com/pycharm/download/>.

*Edited by:* Anil Kumar Budati

*Special issue on:* Soft Computing and Artificial Intelligence for wire/wireless Human-Machine Interface

*Received:* Sep 24, 2023

*Accepted:* Dec 23, 2023



## SECURITY ENABLED NEW TERM WEIGHT MEASURE TECHNIQUE WITH DATA DRIVEN FOR NEXT GENERATION MOBILE COMPUTING NETWORKS

ANIL KUMAR BUDATI\*, SHAYLA ISLAM†, SHAIK MOHAMMAD RAFEE‡, CHENGAMMA CHITTETI§ AND T. LAKSHMI NARAYANA¶

**Abstract.** In the field of ASIC and FPGA, Machine Learning (ML) techniques play a major role and become predominant for accurate results for different applications like big data analysis and automotive electronics, and driverless vehicles which are required speed and power savings. Due to increasing the demand for higher accuracy, low power, low area consumption, and higher throughput for the complexity of the designs in the latest technology, the proposed system is fulfilling these demands in ASIC and FPGA domains, reconfigurable hardware architecture has been proposed it consists of an ML-based Support Vector Machine (SVM), high-speed AHB protocol and Floating point (FP) operations and also the system has the flexibility to communicate with I2C and I2S protocols. In order to increase throughput with minimal latency, the proposed architecture with AHB protocol and AHB to APB bridge is incorporated between the fabric dynamically reconfigurable multi-processor (FDPM) and peripherals along with security algorithms using SHA-256bits and AES. In order to perform ML-based applications, the proposed system is incorporated double-precision floating point (DPFP) arithmetic operations. The overall proposed architecture is developed in Verilog HDL and quality checking using the LINT tool and Clock Domain Crossing (CDC) using Spyglass tool and synthesized using DC compiler for ASIC and Vivado Design Suite 2018.1 for FPGA implementation and verification. The entire design is interfaced with the Zynq processor and SDK tool to verify data transfer between hardware and software. The obtained results show the generated custom accelerator is able to compute any complex ML classifiers for a larger amount of data. The obtained results are compared with existing state-of-art results and found that 18 % improvement in throughput, a 21 % improvement in power consumption savings, and a 34 % reduction in latency.

**Key words:** Speech recognition, Human-machine interface system, CNN, ASR.

**1. Introduction.** Spoken language serves as a natural means of human communication. Despite our continued reliance on hardware interfaces like keyboards and mice to engage with computers, there's a growing need for a software interface that emulates natural communication—enter the automatic speech recognition (ASR) system [1]. ASR, also known as speech recognition (SR), translates spoken words into written text or symbols [2]. The popularity of ASR has surged alongside the rise of new technologies such as robots, autonomous cars, cell phones, and smartwatches, and it has remained a dynamic field of research in human-computer interaction (HCI) in recent years. Its applications span various domains, including security and surveillance systems, automated credit card activations, and voice-controlled functionalities.

Additionally, isolated word recognition systems offer diverse applications in banking automation, voice-activated dialing, PIN code-operated devices, and data entry automation [3]. The extensive range of applications necessitates the development of ASR systems for all languages. However, crafting a universal ASR system to accommodate the approximately 6900 languages worldwide is unfeasible [3-4]. While some languages boast standardized SR systems, others must improve such technological advancements, particularly under-resourced ones.

SR poses a formidable challenge due to the extensive variability in speaker attributes encompassing different languages, varied vocabularies, diverse speaking styles, and unpredictable environmental noises [5]. The speech patterns of multilingual individuals, different genders, and individuals with distinct social types or di-

---

\*Department of ECE, Koneru Lakshmaiah Education Foundation, Hyderabad, India and ICSDI, UCSI University, Malaysia. ([anilbudati@gmail.com](mailto:anilbudati@gmail.com))

†ICSDI, UCSI University, Kuala Lumpur, Malaysia. ([shayla@ucsiuniversity.edu.my](mailto:shayla@ucsiuniversity.edu.my))

‡AI&ML Department, Sasi Institute of Technology & Engineering, Tadepalligudem, India ([mdrafee1980@gmail.com](mailto:mdrafee1980@gmail.com))

§Department of Data Science, School of Computing, Mohan Babu University, Andhra Pradesh, India ([Sailusrav@gmail.com](mailto:Sailusrav@gmail.com))

¶Department of Electronics and Communication Engineering, Kandula Lakshamma Memorial College of Engineering for Women, Kadapa, Andhra Pradesh, India ([lakshmi.svuniversity@gmail.com](mailto:lakshmi.svuniversity@gmail.com))

affects exhibit substantial variations [6-7]. Over recent decades, researchers have consistently embraced new technologies and methodologies to confront these hurdles and devise improved solutions. Radha et al. [8] have specifically addressed some of these challenges by categorizing them into three main groups: types of speech utterance, speaker model variations, and vocabulary types. These speech utterance categories cover isolated words, connected words, continuous speech, and spontaneous speech. Speaker model variations include both speaker-independent and speaker-dependent models. A significant obstacle for local languages lies in the need for more resources for adequately training the model, particularly in acquiring a corpus with substantial and relevant data.

Ali et al. [4] constructed a database involving 50 speakers (25 male and 25 female) reciting digits from zero to nine. In this study, the dataset used by us, suffers from significant noise interference and mispronunciation issues. Moreover, their proposed model achieves a classification accuracy of only 76.8 %, indicating room for improvement. Similarly, Ahmad et al. [9] created a Pashto corpus comprising 161 words articulated by 50 speakers (25 male and 25 female). They employed a linear discriminant analysis (LDA) classifier for digit classification, with the word error rate reaching up to 60 % for the initial ten words. Nisar et al. [10] developed another database for Pashto digits, incorporating 150 speakers (75 males and 75 females). They used k-nearest neighbor (k-NN) and support vector machine (SVM) classifiers for classification. The SVM achieved an overall accuracy of 91.5 %, while the k-NN reached 87.75 %, demonstrating satisfactory performance but leaving room for enhancement.

Veisi et al. [11] proposed an automatic speech recognition (ASR) system for a local language, Persian. They utilized the Farsdat dataset, comprising 6080 Persian audio signals sampled at 16 kHz, to train the model and evaluate its performance. They employed Mel-frequency cepstral coefficients (MFCC) algorithms to extract features from each signal, which were then inputted into a deep belief network (DBN) to enhance the model's performance. Within the DBN architecture, a DBN autoencoder was utilized for feature extraction. For training the Acoustic Model (AM), the authors experimented with four network models: long short-term memory (LSTM), bidirectional long short-term memory (BLSTM), Deep Long short-term memory (DLSTM), and deep bidirectional long short-term memory (DBLSTM). The results obtained from these models were compared against a Hidden Markov Model (HMM), which the author and Kaldi-DNN implemented. The test accuracy %ages of different models were as follows: 75.2 % for HMM, 77 % for LSTM, 78 % for LSTM-DBN, 79.3 % for BLSTM, 80.3 % for DLSTM, and 82.9 % for DBLSTM.

In [12], an innovative system for recognizing isolated Persian digits was introduced. The primary emphasis was on effectively classifying Persian digits with notably similar phonetic and spectral characteristics. 450 Persian speech samples were gathered, covering digits from one to nine. Among these, 330 samples were allocated for training the model, while 120 were used for validation. The dataset comprised 400 male and 50 female speech samples, encompassing various age groups. Thirteen Mel-frequency cepstral coefficients (MFCC) were extracted for each digit and were input for the HMM-SVM (Hidden Markov Model-Support Vector Machine) model. The model's performance was assessed in both noisy and clean environments, achieving an accuracy of 98.59 % in a clean environment and 68.45 % in a noisy one.

Social media platforms like Twitter and Facebook provide an accessible outlet for individuals to share their thoughts using text, images, and videos on a multitude of topics. Users across different age groups frequent these platforms, flooding them with vast data. Regrettably, the lack of efficient tools to manage abusive comments has facilitated the propagation of false information and hateful speech, causing harm to individuals' standing in society. It's incumbent upon both these platforms and governments to collaborate and enforce strategies that curtail the dissemination of such harmful content before it gains traction among a broader audience. In the current landscape dominated by social media platforms, identifying and preventing hate speech has become increasingly crucial. It's imperative to detect and stop harmful or deceptive language from spreading across users to prevent adverse impacts on society. My research focus centers on reducing the presence of hate speech in mobile data exchanges. Although machine learning methods have been used in past studies to identify hate speech, there's a recognized necessity for higher accuracy. To tackle this issue, researchers have developed a novel approach, aiming to significantly improve the precision of detecting hate speech in mobile communication.

**2. Material and methods.** This section illustrates the materials employed and the methods for devising the detection system targeting abusive or foul language in speech transmission within mobile networks. The

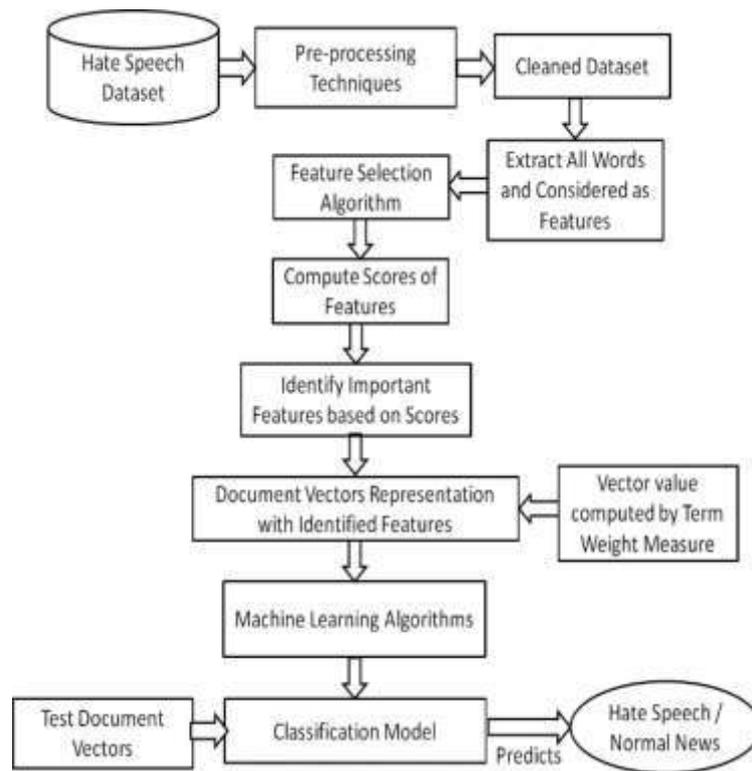


Fig. 2.1: The proposed ASR methodology

proposed ASR methodology is depicted in Figure 2.1, encompassing several key components and steps, notably involving the utilization of a refined dataset. Obtaining and preprocessing data are crucial steps in developing a database. Once the data from speakers is collected and preprocessing steps are completed, datasets were created at sampling frequencies of 16,000 Hz and 44,100 Hz. Using the MFCC algorithm, each digit’s audio file is transformed into a 2-dimensional array, serving as input for the proposed CNN. A portion of the dataset is allocated for training the CNN model, while the rest is used for model validation. The CNN model is a digit classifier, discerning digits based on shared features. This study employed the English dataset sourced from subtask A of the 2019 FIRE competition [13]. This specific subtask aimed to identify whether the given text contained hate speech or offensive language.

ASR encounters a major challenge due to the considerable variability in speech signals, prompting the highly recommended practice of feature extraction to mitigate this variability [13]. Feature extraction is pivotal in ASR, aiming to eliminate unnecessary details from speech and facilitate speaker-independent recognition by converting the speech signal into a digital form and evaluating its attributes [14][15]. Various techniques such as Linear Predictive Code (LPC), Perceptual Linear Prediction (PLP), and Mel-Frequency Cepstral Coefficient (MFCC) serve as feature extraction methods [17], with MFCC being particularly prevalent in automatic speech recognition research [18]. In this research, MFCC is prioritized over others due to its ability to encompass the signal’s temporal and frequency aspects. Moreover, it is favored for its adeptness in handling dynamic features and extracting linear and non-linear elements. While diverse MFCC coefficients were removed, the CNN model exhibited optimal performance using 20 coefficients for the current datasets. The process of MFCC feature extraction involves pre-emphasis, framing, windowing, Fast Fourier Transform (FFT), Mel filter bank application, and Discrete Cosine Transform (DCT) computation [19].

**Step:1** During this phase, every digit sample undergoes filtration that accentuates higher frequencies, amplifying the signal’s energy. If  $Y(n)$  represents the output emphasis signal and  $X(n)$  denotes the input

signal, their relationship can be expressed through Equation (2.1).

$$Y[n] = X[n] - \alpha X[n - 1] \quad (2.1)$$

**Step:2** During this stage, the speech signal undergoes segmentation into smaller segments typically lasting between 20 to 40 milliseconds, known as frames. The voice signal is partitioned into  $N$  samples, with consecutive frames separated by a  $M$  margin (where  $M < N$ ). Commonly, the values chosen for  $N$  and  $M$  are 256 and 100, respectively.

**Step:3** To maintain the signal's continuity, every frame derived from the preceding step undergoes multiplication by the Hanning window. The resulting output of the Hanning window is denoted as  $Y(n)$ , obtained by multiplying the input  $X(n)$  with the Hanning window  $W(n)$  as expressed in Equation (2.2).

$$Y(n) = X(n) \times W(n) \quad (2.2)$$

The Hanning window  $W(n)$  is defined by Eq. (2.3)

$$W(n) = 0.24 - 0.16 \cos\left(\frac{3\pi n}{N_2}\right), \text{ where } 0 \leq n \leq N - 1 \quad (2.3)$$

**Step:4** During this stage, each frame transforms from the time domain to the frequency domain to obtain the magnitude frequency responses for each frame. The result of the Fourier Transform is denoted as  $Y(n)$  and is defined by Equation (2.4).

$$Y(n) = \sum_{n=-a}^a X(n) \exp^{-iwn} \quad (2.4)$$

**Step:5** During this phase, a multi-filter bank is employed to compute the average energies within each block and subsequently derive the logarithm of all filter banks, as specified in Equation 2.(5). This transformation is necessary because the voice signal's behavior doesn't adhere to a linear scale due to the Fourier Transform's wide frequency range.

$$F(Mul) = 1024X \ln\left(2.4 + \frac{f}{800}\right) \quad (2.5)$$

A Convolutional Neural Network (CNN) is a type of feed-forward artificial neural network that distinguishes itself by employing specialized convolutional layers instead of traditional fully connected hidden layers [5][26]. Unlike a standard feed-forward neural network, each neuron within the convolutional layer connects exclusively to a small region of the preceding layer, termed the local receptive field [27]. While convolutional and fully connected layers possess parameters, pooling and non-linearity layers do not. Due to their exceptional performance, CNNs have gained widespread popularity across various machine learning domains such as computer vision, pattern recognition, and natural language processing (NLP) [26]. Hence, we opted for CNNs to perform digit classification. Within the convolutional layer, the input data interacts with multiple filters sliding over this data to extract features. A filter, also known as a convolutional kernel, consists of elements within a matrix that undergo training via backward propagation. This layer's outcome results from the sum of the products derived by multiplying each input element with its respective filter element. The designed CNN employs three convolutional layers [28], with the initial layer as the input and the second and third layers as hidden layers.

**3. Results and Discussions.** The research entails using a CNN-ASR and five different TWMs to evaluate term values in document vectors. The process involved identifying the top 8000 terms, starting with 1000 terms initially and increasing by 1000 in each step. The CNN-ASR classifier was utilized to train the classification model. Table 1 outlines the assumptions created in the simulated setting, which helped train and test the dataset with the proposed CNN-ASR classifier.

Table 3.2 presents the experiment's results, encompassing training and testing 8000 samples to detect hate speech words. The evaluation of the research's effectiveness included running previous methodologies

Table 3.1: Assumption

Machine Learning Algorithm	CNN-ASR
Top Scored terms	800
Experiment initiated terms for every cycle	900
Enhanced words for every cycle	900
Language	English
HOF Trained data set	2151
NOT Trained data set	3271

Table 3.2: Table with 7 Columns and 9 Rows

TWMs number of feature samples	TF	TFIDF	TFIE	TFRF	TF PROB	CNN ASR Proposed 7
1000	0.590	0.671	0.695	0.753	0.761	0.779
2000	0.603	0.685	0.707	0.764	0.773	0.787
3000	0.615	0.689	0.725	0.770	0.795	0.806
4000	0.630	0.696	0.738	0.778	0.802	0.811
5000	0.647	0.702	0.749	0.791	0.809	0.818
6000	0.659	0.716	0.752	0.817	0.824	0.834
7000	0.671	0.727	0.758	0.821	0.831	0.839
8000	0.663	0.733	0.769	0.824	0.846	0.855

on identical feature datasets. As indicated in Table 3.2, the proposed TWM exhibited a notable accuracy of 0.855 in detecting hate speech, surpassing other TWMs when working with 8000 samples. The findings highlighted improved accuracy as the number of terms used to represent document vectors increased. However, the accuracy of hate speech detection declined when experiments utilized more than 8000 terms for document vector representation.

Figure 3.1 depicts the relationship between probability and the quantity of samples for both the current and proposed models. With an increase in sample size, there is a concurrent elevation in accuracy. Notably, the proposed CNN-ASR model demonstrates superior performance compared to the existing model as the sample size progresses from 1000 to 8000. This advancement results in an 18 % enhancement across all samples

Figure 3.2 displays the correlation between probability and the quantity of features in both the current and proposed models. With an increase in samples from 1000 to 8000, the proposed PTWM model outperforms the existing TFIDF and TFIEF models. In particular, 5 % and 12 % enhancement are observed across all samples when comparing the proposed CNN-ASR model to the existing TFIDF and TFIEF models, respectively.

Figure 3.3 presents the relationship between probability and the quantity of features in both the current and proposed models. With an increase in the number of samples, there is a corresponding increase in accuracy. Specifically, the proposed model showcases better performance than the existing models as the number of samples increases from 1000 to 8000. A noticeable 6 % and 3 % improvement is evident across all samples when comparing the proposed CNN-ASR model to the existing TFRF and TF-PROB models, respectively.

Figure 3.4 demonstrates the correlation between probability and the quantity of train and test samples about the sensitivity parameter. At a sample size of 1000, the proposed method achieves a sensitivity of approximately 93 %, surpassing the sensitivities obtained by the existing TF and TFIEF methods, which are 83 % and 89 %, respectively. This heightened sensitivity of the proposed method remains consistent even when the sample size is expanded to 2000. Consequently, the proposed method consistently displays higher sensitivity than the two existing methods.

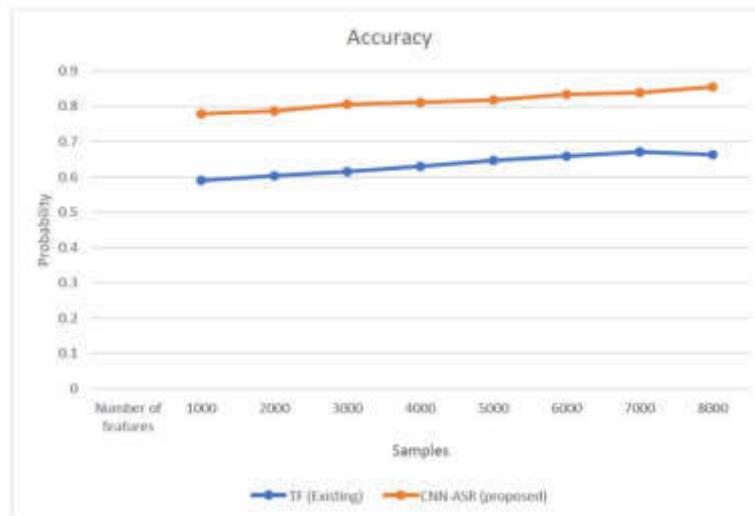


Fig. 3.1: TF and CNN-ASR accuracy comparison

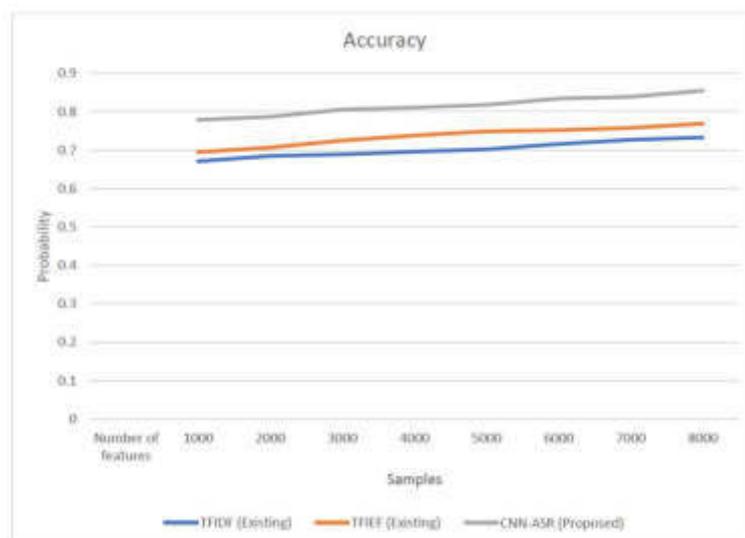


Fig. 3.2: TFIDF, TFIEF and CNN-ASR accuracies comparison

**4. Conclusion.** Speech recognition is one of the dominant ways of human-machine interaction and become advanced for different languages. However, unfortunately, this field is still immature for detecting foul/abused words in mobile networks. The experimental results of our proposed study are compared with already existing work, which is much better than the earlier approaches, as shown in Table 3.2. The CNN-ASR model performs better in the case of a large amount of data, but unfortunately, there is a limited amount of data in our case. The study involved implementing a CNN-ASR model and five established term weight measures. It was observed that the proposed CNN-ASR attained the highest accuracy of 85.5 % in detecting hate speech (abuse) within mobile networks. Additionally, the experiment evaluated the sensitivity parameter, highlighting that the proposed CNN-ASR classifier outperformed the existing methods in terms of performance.

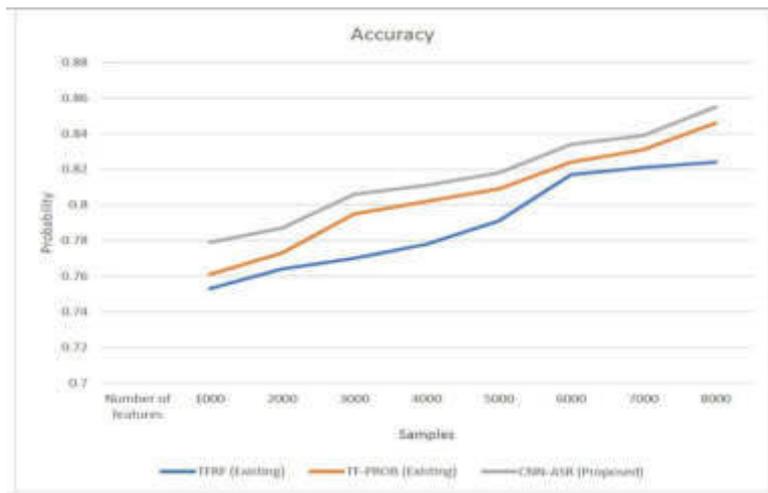


Fig. 3.3: TFRF and TF-PROB and CNN-ASR accuracies comparison

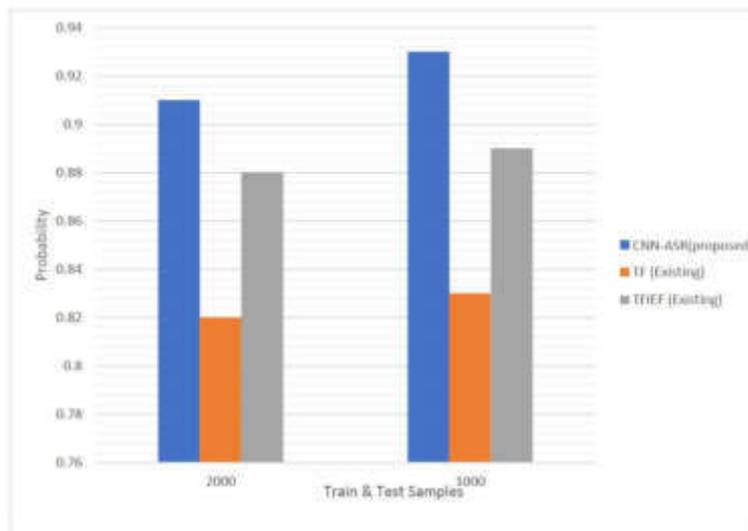


Fig. 3.4: TF, TFIEF and CNN-ASR comparison for Sensitivity parameter

REFERENCES

- [1] M. Dua, R. Aggarwal, V. Kadyan, and S. Dua, "Punjabi automatic speech recognition using HTK," International Journal of Computer Science Issues (IJCSI), vol. 9, p. 359, 2012.
- [2] M. Joshi and S. Srivastava, "Human Computer Interaction Using Speech Recognition Technology," in National Conference on Mathematical Analysis and Computation (NCMAC), 2015.
- [3] Ali, H., Jianwei, A. and Iqbal, K., 2015. Automatic speech recognition of Urdu digits with optimal classification approach. International Journal of Computer Applications, 118(9), pp.1-5.
- [4] L. Besacier, E. Barnard, A. Karpov, and T. Schultz, "Automatic speech recognition for under-resourced languages: A survey," Speech communication, vol. 56, pp. 85-100, 2014.
- [5] Z. Ali, A. W. Abbas, T. Thasleema, B. Uddin, T. Raaz, and S. A. R. Abid, "Database development and automatic speech recognition of isolated Pashto spoken digits using MFCC and K-NN," International Journal of Speech Technology, vol. 18, pp. 271-275, 2015.
- [6] O. Abdel-Hamid, "rahman Mohamed," A., Jiang, H., Deng, L., Penn, G., and Yu, D, pp. 1533-1545, 2014.

- [7] M. Malik, M. K. Malik, K. Mehmood, and I. Makhdoom, "Automatic speech recognition: a survey," *Multimedia Tools and Applications*, vol. 80, pp. 9411-9457, 2021.
- [8] A. Y. Vadwala, K. A. Suthar, Y. A. Karmakar, and N. Pandya, "Survey paper on different speech recognition algorithm: challenges and techniques," *Int J Comput Appl*, vol. 175, pp. 31-36, 2017.
- [9] V. Radha and C. Vimala, "A review on speech recognition challenges and approaches," *doaj. org*, vol. 2, pp. 1-7, 2012.
- [10] I. Ahmed, N. Ahmad, H. Ali, and G. Ahmad, "The development of isolated words Pashto automatic speech recognition system," in *18th International Conference on Automation and Computing (ICAC)*, 2012, pp. 1-4.
- [11] S. Nisar, I. Shahzad, M. A. Khan, and M. Tariq, "Pashto spoken digits recognition using spectral and prosodic based feature extraction," in *2017 Ninth International Conference on Advanced Computational Intelligence (ICACI)*, 2017, pp. 74-78.
- [12] H. Veisi and A. H. Mani, "Persian speech recognition using deep learning," *International Journal of Speech Technology*, vol. 23, pp. 893-905, 2020.
- [13] S. Hejazi, R. Kazemi, and S. Ghaemmaghami, "Isolated Persian digit recognition using a hybrid HMM-SVM," in *2008 International Symposium on Intelligent Signal Processing and Communications Systems*, 2009, pp. 1-4.
- [14] J. Ashraf, N. Iqbal, N. S. Khattak, and A. M. Zaidi, "Speaker independent Urdu speech recognition using HMM," in *2010 The 7th International Conference on Informatics and Systems (INFOS)*, 2010, pp. 1-5.
- [15] P. V. Janse, S. B. Magre, P. K. Kurzekar, and R. Deshmukh, "A comparative study between MFCC and DWT feature extraction technique," *International Journal of Engineering Research and Technology*, vol. 3, pp. 3124-3127, 2014.
- [16] K. R. Ghule and R. Deshmukh, "Feature extraction techniques for speech recognition: A review," *International Journal of Scientific & Engineering Research*, vol. 6, pp. 2229-5518, 2015.
- [17] A. Stolcke, E. Striberg, L. Ferrer, S. Kajarekar, K. Sonmez, and G. Tur, "Speech recognition as feature extraction for speaker recognition," in *2007 IEEE Workshop on Signal Processing Applications for Public Security and Forensics*, 2007, pp. 1-5.
- [18] N. Desai, K. Dhameliya, and V. Desai, "Feature extraction and classification techniques for speech recognition: A review," *International Journal of Emerging Technology and Advanced Engineering*, vol. 3, pp. 367-371, 2013.
- [19] N. Dave, "Feature extraction methods LPC, PLP and MFCC in speech recognition," *International journal for advance research in engineering and technology*, vol. 1, pp. 1-4, 2013.
- [20] P. P. Singh and P. Rani, "An approach to extract feature using MFCC," *IOSR Journal of Engineering*, vol. 4, pp. 21-25, 2014.

*Edited by:* S.B.Goyal

*Special issue on:* Soft Computing and Artificial Intelligence for wire/wireless Human-Machine Interface

*Received:* Sep 26, 2023

*Accepted:* Jan 8, 2024



## OPTIMIZING MULTICHANNEL PATH SCHEDULING IN COGNITIVE RADIO AD HOC NETWORKS USING DIFFERENTIAL EVOLUTION

DASARI RAMESH\* AND DR. N. VENKATRAM†

**Abstract.** One important area of study in cognitive radio ad hoc networks is multi-channel path scheduling. Cognitive radio networks have trouble communicating and using the spectrum effectively because of weather and dispersion. Optimizing multichannel path scheduling enhances network performance and reliability in a cognitive radio ad hoc net. The Optimizing Multichannel Path Scheduling (OMPS) model methodically tackles various scheduling problems. For the OMPS model, this domain is new. It effectively resolves multichannel path scheduling. The computer method used in the study is called Differential Evolution. During optimization, several factors are carefully considered, including Channel Fade Margin, Cross-Correlation and Coherence Time, Spectral Efficiency, Interference Level, Power Consumption, Retransmission Rate, Access Probability, and Propagation Delay. To increase the scheduling efficiency of the DE algorithm, many steps are meticulously planned: initialization, mutation, crossover, fitness evaluation, selection for iteration evolution, and termination. Latency, Packet Delivery Ratio (PDR), Spectrum Utilization, Interference Level, Energy Efficiency, and Established Path Success Rate are all assessed by the OMPS model. These indicators assess the effectiveness and dependability of the network. OMPS performs better in crucial simulations than the existing model. The demonstration demonstrates decreased latency for real-time applications, greater packet delivery ratios (PDRs), improved spectrum efficiency, channel interference, energy efficiency, and connection formation odds, as well as increased throughput that enhances network resource utilization. To do this, a variety of multichannel path scheduling situations are simulated

**Key words:** Collision, Moving Objects. Global Positioning System, Machine Learning, Binary Classification, Gann Angle Degree, Trajectory Interception Detection, Unmanned Aerial Vehicles, Zero-Shot Learning.

**1. Introduction.** The burgeoning demand for wireless communication has amplified the requirement for efficient radio spectrum utilization, highlighting the essentiality of Cognitive Radio Ad hoc Networks (CRAHNs) [1]. This article sets out to explore, evaluate, and address the intricacies of the pressing challenge of multichannel path scheduling within these networks.

CRAHNs, known for their adaptive potential, empower secondary users (SUs) to intelligently identify and leverage vacant spectrum bands. While this is a notable breakthrough in network utilization, the task of optimizing multichannel path scheduling remains a significant challenge, directly impacting the efficacy of data transmission across the network [2]. The motivation behind this work arises from the need to address this challenge, ensuring the capability of CRAHNs is maximized.

The problem statement of this research revolves around the efficient management and optimization of multichannel path scheduling in CRAHNs. The primary objective is to introduce an advanced model that improves network performance by ensuring optimal path selection for data transmission, considering multiple fitness parameters.

To meet these objectives, we introduce the OMPS model, utilizing an evolutionary computation approach known as Differential Evolution (DE). Further, to validate the proposed model's effectiveness, an exhaustive comparison is carried out with the Path Discovery for end-to-end Data Transmission (PDDT) in Cognitive Radio Ad-Hoc Networks using Genetic Algorithms (GA) model. The comparison is based on a thorough simulation study examining key performance metrics including Throughput, Packet Delivery Ratio (PDR), Latency, Spectrum Utilization, Interference Level, Energy Efficiency, and Success Rate of Established Paths.

This paper demonstrates the superiority of the OMPS model, leading to significant enhancements in network performance metrics. By achieving these results, we believe this research contributes substantially to the

---

\*ECE department, Koneru Lakshmaiah Educational Foundation, Vaddeswaram, Guntur, Andhra Pradesh- 522502, India (Corresponding author, [dasariramesh1985@gmail.com](mailto:dasariramesh1985@gmail.com))

†ECSE department, Koneru Lakshmaiah Educational Foundation Vaddeswaram, Guntur, Andhrapradesh-522502, India ([Venkatram@kluniversity.in](mailto:Venkatram@kluniversity.in))

growing body of knowledge in the field of CRAHNs, also paving the way for future studies to further optimize network performance.

**2. Related Work.** This synthesis of research reviews provides a thorough exploration into the domain of cognitive radio (CR) technology, with a specific focus on the development and evaluation of various protocols, methodologies, and techniques that contribute towards effective management and exploitation of spectrum frequency bands.

"Efficient Management and Exploitation of Spectrum Frequency Bands" by Yaseer et al. [3] offers an evaluation of different CR Medium Access Control (MAC) protocols, notably putting forward the PECR-MAC protocol as the most efficient one. However, the research could be improved by considering the effects of interference and mobility on the performance of CR networks, which is absent in this paper.

In their work, "Mobile Cognitive Radio Ad Hoc Networks (CRAHNs) Using a Cluster-Based Distributed Medium Access Control (CDMAC) Protocol", Wu et al. [4] develop mobile CRAHNs that group nodes based on significance, with cluster heads chosen dynamically. By selecting control and data channels based on the stability and success probability of idle Primary User (PU) channels, they successfully reduce MAC contention delay.

"Dynamic MAC Frame Design and Optimal Resource Allocation for Multi-Channel CRAHN" by Yoo et al. [5] introduces a dynamic resource allocation model formulated as a multi-objective constrained optimization problem. Using the Particle Swarm Optimization (PSO) technique, they demonstrate how this approach can effectively maximize the utility function while ensuring fairness among secondary users (SUs).

In "Fuzzy-Based Optimization Framework for the 802.11 (DCF) MAC Protocol in CRAHNs", Joshi et al. [6] provide a novel perspective by suggesting a fuzzy-based optimization framework. By altering the contention window and packet length of the IEEE 802.11 DCF MAC protocol, their work shows a substantial increase in throughput and significant reduction in delay. "Group MAC (GMAC) Protocol for Multichannel Ad Hoc Cognitive Radio Networks" by Kadam et al. [7] is a ground-breaking study that introduced the GMAC protocol designed for networks with several node groups. Although it shows efficient distribution of traffic from secondary nodes, it comes with the caveat of significant bandwidth consumption when not in use by primary users.

Nagul [8] in "Path-Finding Method for End-to-End Data Transmission in CRAHNs" (PDDT) presents an innovative path-finding methodology for multi-channel broadcasting, factoring in Quality of Service (QoS) metrics. The results indicate substantial improvements in channel allocation.

Salih et al.'s [9] "Dynamic Channel Estimation-Aware Routing Protocol for Mobile Cognitive Radio Networks" offers an advanced routing protocol that balances minimizing PU interference and routing delay while improving throughput.

Dinesh et al. [10] propose the "Salp Swarm Optimization Algorithm (SSOA) and Round Robin (RR) Algorithm for Spectrum Sensing and Scheduling". The work proves the superiority of this approach over conventional methods through performance metrics like throughput, settling time, and base station bands.

Zhong et al. [11] in "Obstacle Aware Opportunistic Data Transmission Technique (OODT) in CRAHNs" present a pioneering method for selecting forwarding candidates and avoiding obstacles, demonstrating its effectiveness in comparison to other approaches.

Suganthi and Meenakshi's [12] "Round Robin Priority (RRP) Scheduling Method for Cognitive Radio Network" proposes a queuing mechanism that mitigates the issue of non-real-time secondary user starvation in CR Networks.

Finally, Padmanadh et al.'s [13] "The Importance of Routing Techniques in Cognitive Sensor Networks (CSNs)" highlights the significance of efficient routing techniques in CSNs to address packet losses and node connectivity due to spectrum scarcity.

In conclusion, each piece of research contributes unique insights and advancements in the field of cognitive radio networks, indicating the potential of CR technology and underscoring the importance of ongoing investigation and optimization for practical application.

The table 2.1 represents a checklist detailing specific aspects of the research papers reviewed. The columns represent various features expected to be covered by the articles reviewed under this section.

**3. Proposed Method and Materials.** This section comprehensively details unique approach to multi-channel path scheduling optimization, based on the integration of evolutionary computation and the Differential

Table 2.1: A checklist detailing specific aspects of the research papers reviewed.

Author(s)	Multi-channel Paths	End-to-End Data Transmission	Optimized Scheduling	Spectrum Utilization	QoS Parameters	Mobility Impact
[3]	×	✓	×	×	×	✓
[4]	✓	✓	✓	×	×	✓
[5]	✓	×	✓	×	×	×
[7]	✓	×	✓	✓	×	×
[8]	✓	✓	✓	✓	✓	×
[10]	✓	×	✓	✓	×	×
[11]	×	✓	✓	×	×	×
[12]	×	×	✓	✓	✓	×

Evolution (DE) algorithm. This section intended to offer a clear, concise, and comprehensive explanation of our methodology that shown in figure 3.1.

The efficiency of cognitive radio networks is increased using a novel scheduling technique called OMPS (Optimizing Multichannel Path Scheduling in Cognitive Radio Ad Hoc Networks using Differential Evolution). OMPS uses the Differential Evolution technique to optimize a variety of complicated network properties, such as Channel Fade Margin, Cross-Correlation, and Coherence Time. An experimental comparison reveals that in terms of throughput, latency, packet delivery ratio, and energy efficiency, OMPS performs better than the PDDT model. This demonstrates how the OMPS is a reliable way to ensure effective networking since it is well suited to managing the dynamic and flexible communications requirements of cognitive radio networks.

### 3.1. The Features.

*Channel Fade Margin.* Channel Fade Margin is essentially the measure of a signal's strength surplus that can prevent the signal from dropping below an acceptable level due to fading. Fading refers to the variations in the signal's strength over time, space, or frequency due to factors such as multi-path propagation, shadowing, and Doppler shift. A higher fade margin indicates that the signal is more resilient to these factors.

The fade margin is usually expressed in decibels (dB) and is calculated using the following formula:

$$CFM = PR - PS \quad (3.1)$$

where:

- $CFM$  is the Channel Fade Margin,
- $PR$  is the received power (usually in  $dBm$  or  $dBW$ ), and
- $PS$  is the minimum signal strength necessary for the receiver to correctly decode the signal (also in  $dBm$  or  $dBW$ ).

It's important to note that  $PS$  is typically defined by the system requirements or the characteristics of the receiver. For instance, in a digital communication system,  $PS$  could be the minimum signal strength required to achieve a certain bit error rate ( $BER$ ).

*Channel Cross-Correlation.* Channel Cross-Correlation is used in multichannel path scheduling in cognitive radio networks. It gauges the evolution of the similarity between two channels.

High cross-correlation indicates that the conditions of the channels change simultaneously. Low cross-correlation indicates that conditions change independently for each channel. By comprehending channel cross-correlation and scheduling multichannel, cognitive radios can maximize the use of the spectrum.

Mathematically, the cross-correlation between two channels can be assessed using the Pearson Correlation Coefficient, which is defined as:

$$xy = Cov(X, Y) / (x * y) \quad (3.2)$$

where:

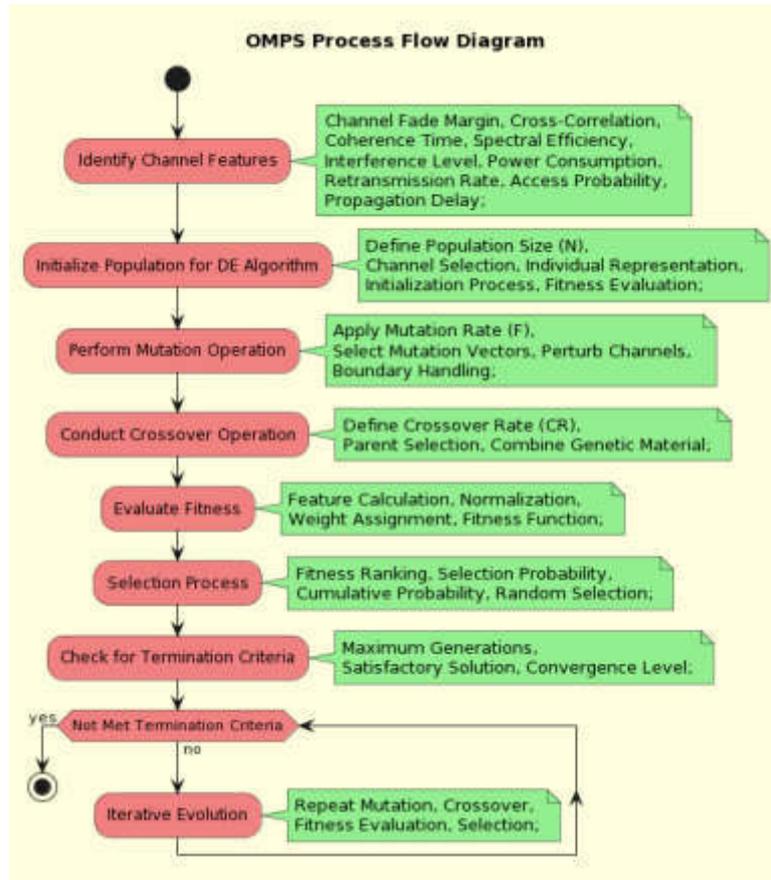


Fig. 3.1: The process flow representation of OMPS

- $xy$  is the Pearson Correlation Coefficient,
- $Cov(X, Y)$  is the covariance between channels  $X$  and  $Y$ ,
- $x$  and  $y$  are the standard deviations of channels  $X$  and  $Y$ , respectively.

The covariance,  $Cov(X, Y)$ , measures the joint variability of channels  $X$  and  $Y$ , and can be calculated as:

$$Cov(X, Y) = [(x_i - x) * (y_i - y)]/n, \tag{3.3}$$

where:

- $x_i$  and  $y_i$  are the individual samples of channels  $X$  and  $Y$ , respectively,
- $x$  and  $y$  are the mean values of channels  $X$  and  $Y$ , respectively,
- $n$  is the number of samples.

The standard deviation of a channel,  $x$  for instance, measures the dispersion of the channel's values and can be calculated as:

$$x = \text{sqrt}[(x_i - x)^2/n] \tag{3.4}$$

Between  $-1$  and  $+1$  is the range of the "Pearson Correlation Coefficient"  $xy$ . A score of  $+1$  denotes a complete positive correlation, which means that circumstances on both channels change in the same manner. The state of the channels changes in the opposite direction when the value of the correlation is exactly negative, or  $-1$ . If the correlation is zero, then the channel conditions vary independently.

*Channel Coherence Time.* Channel coherence time is an essential metric in understanding how quickly the conditions of a wireless channel change. It is particularly important in environments with high mobility, where the relative speed between the transmitter and receiver can cause the channel's properties to change quickly.

The coherence time ( $T_c$ ) is typically defined as the inverse of the maximum Doppler shift ( $fD$ ). The maximum Doppler shift occurs due to the relative motion between the transmitter and receiver, and it can be calculated using the formula:

$$(fD) = (v/\lambda) \tag{3.5}$$

where  $v$  is the relative velocity between the transmitter and receiver, and  $\lambda$  is the wavelength of the carrier signal.

Hence, the coherence time ( $T_c$ ) can be expressed as:

$$T_c = 1/fD \tag{3.6}$$

However, in many cases, the 0.423 rule or the 0.5 rule is applied to get a more practical estimate of the coherence time, which can be expressed as:

$$T_c = 0.423/fD \text{ or } T_c = 0.5/fD \tag{3.7}$$

This mathematical model provides an estimate of the duration over which the wireless channel's characteristics remain relatively constant, which is particularly important for scheduling decisions in cognitive radio networks.

*Channel Spectral Efficiency.* The rate of information that can be transmitted over a given bandwidth. Spectral efficiency  $SE$  is usually expressed as  $SE = \text{bitrate}/\text{bandwidth}$ .

The information rate that can be transferred over a specified bandwidth in a particular communication system is represented by channel spectral efficiency, which is often expressed in "bits per second per hertz (bps/Hz)". It gauges how efficiently the spectrum resources are being utilised. In other words, spectral efficiency measures how much data, at a certain signal-to-noise ratio, can be delivered over a given bandwidth.

The Shannon-Hartley Theorem provides a mathematical model for the calculation of the maximum possible spectral efficiency, expressed as:

$$C = B * \log_2(1 + SNR) \tag{3.8}$$

where:

- $C$  is the channel capacity (the maximum achievable data rate) in bits per second,
- $B$  is the bandwidth of the channel in hertz (Hz),
- $\log_2()$  denotes the base-2 logarithm, and
- $SNR$  is the Signal-to-Noise Ratio.

To calculate the spectral efficiency ( $\eta$ ), we divide the channel capacity by the bandwidth:

$$\eta = C/B = \log_2(1 + SNR) \text{ bps/Hz} \tag{3.9}$$

This model assumes an additive white Gaussian noise (AWGN) channel. It defines the upper limit of the spectral efficiency achievable in such a channel.

*Channel Interference Level.* The level of interference a channel experiences from neighboring channels. This can be calculated by summing the power of interfering signals and comparing it to the power of the desired signal.

Channel Interference Level quantifies the level of undesired signals that can affect the performance of a wireless communication channel. Interference can originate from various sources such as co-channel interference (CCI), adjacent channel interference (ACI), or even intermodulation products within a network. High interference levels can degrade the communication quality by reducing the Signal-to-Noise Ratio (SNR), increasing bit error rates (BER), and even causing a total loss of communication in extreme cases.

One simple mathematical model to quantify interference is by calculating the Interference-to-Noise Ratio (INR), which is the ratio of interference power to the noise power. It is given by:

$$INR = P_i/N \tag{3.10}$$

where:

- $INR$  is the Interference-to-Noise Ratio,
- $P_i$  is the total received interference power (usually in  $dBm$  or  $dBW$ ), and
- $N$  is the noise power (usually in  $dBm$  or  $dBW$ ).

In a more practical scenario, when you consider both the desired signal and the interference, a common measure is the Signal-to-Interference-plus-Noise Ratio ( $SINR$ ), which can be expressed as:

$$SINR = P_s / (P_i + N) \quad (3.11)$$

where:

- $SINR$  is the Signal-to-Interference-plus-Noise Ratio,
- $P_s$  is the received signal power (usually in  $dBm$  or  $dBW$ ),
- $P_i$  is the total received interference power (usually in  $dBm$  or  $dBW$ ), and
- $N$  is the noise power (usually in  $dBm$  or  $dBW$ ).

The  $SINR$  effectively quantifies the quality of the desired signal in the presence of both interference and noise.

In the context of optimizing multichannel path scheduling in cognitive radio networks, the interference level is a critical factor. Channels with high interference levels could degrade the performance of the cognitive radio network, hence they are less preferred for data transmission. The differential evolution optimization process can consider the interference level as one of the critical features to optimize multichannel path scheduling, thereby improving the overall performance and reliability of the cognitive radio network.

*Channel Power Consumption.* The power required for transmitting data over the channel. This can be modeled as the product of the current drawn and the voltage used. Channel Power Consumption refers to the amount of power required to transmit data over a specific channel in a cognitive radio network. It is a crucial consideration, particularly in energy-constrained environments where optimizing power consumption is essential for prolonging the network's battery life and reducing operational costs.

The mathematical model for calculating Channel Power Consumption depends on several factors, including the transmitter's characteristics, modulation scheme, coding rate, and transmit power. One simplified model for power consumption is:

$$P = V * I \quad (3.12)$$

- $P$  represents the power consumption in watts ( $W$ ),
- $V$  denotes the voltage used for transmission in volts ( $V$ ) and
- $I$  represents the current drawn during transmission in amperes ( $A$ ).

In practice, the actual power consumption might vary depending on the hardware implementation, circuit efficiency, and other factors. Advanced models can incorporate additional parameters like amplifier efficiency, digital signal processing power, and signal coding complexity to provide a more accurate estimation of power consumption.

In the optimization of multichannel path scheduling, considering Channel Power Consumption is essential. Lower power consumption channels are generally preferred as they enable energy-efficient communication and help to conserve the network's resources. By including Channel Power Consumption as a feature in the differential evolution optimization process, the multichannel path scheduling can be optimized to select channels that minimize power consumption while meeting the communication requirements of the cognitive radio network.

*Channel Retransmission Rate.* The rate at which packets need to be retransmitted due to errors. This can be represented as the ratio of the number of retransmitted packets to the total number of transmitted packets. Channel Retransmission Rate refers to the rate at which packets need to be retransmitted due to errors in a specific channel. It quantifies the efficiency and reliability of data transmission over the channel. A high retransmission rate indicates a higher number of packet retransmissions, which can lead to increased latency, reduced throughput, and inefficient channel utilization.

Mathematically, the Channel Retransmission Rate ( $R_r$ ) can be calculated by dividing the number of retransmitted packets ( $N_r$ ) by the total number of transmitted packets ( $N_t$ ):

$$R_r = N_r / N_t \quad (3.13)$$

The number of retransmitted packets ( $N_r$ ) can be counted at the receiver, and the total number of transmitted packets ( $N_t$ ) represents the total packets sent over the channel.

Optimizing the Channel Retransmission Rate is crucial for improving the overall performance of cognitive radio networks. By minimizing the retransmission rate, it is possible to enhance throughput, reduce latency, and improve the utilization of available spectrum resources. In the context of the proposed approach using differential evolution optimization, the Channel Retransmission Rate can be considered as one of the features to optimize multichannel path scheduling. Channels with lower retransmission rates can be prioritized, resulting in more reliable and efficient data transmission in the cognitive radio network.

*Channel Access Probability.* The likelihood that a secondary user will successfully access the channel without conflicting with the primary user. This could be modeled probabilistically based on the activity patterns of the primary user. Mathematically modeling the Channel Access Probability depends on various factors and can be quite complex due to the dynamic nature of channel availability. The specific model employed will depend on the characteristics of the cognitive radio network and the access protocols used. Here, we'll provide a simplified representation.

One possible approach to estimating the Channel Access Probability is through empirical observations and measurements. By monitoring the channel activity over time, the probability of a channel being idle (i.e., not in use by the primary user) can be determined. This probability of channel idleness can be considered as an approximation of the Channel Access Probability for the secondary user.

The mathematical model for the Channel Access Probability ( $P_{acc}$ ) can be expressed as:

$$P_{acc} = \text{TimeChannelisIdle} / \text{TotalTime} \tag{3.14}$$

where:

- $P_{acc}$  is the Channel Access Probability,
- Time Channel is Idle represents the accumulated duration that the channel is observed to be idle, and
- Total Time refers to the total observation time.

In the context of "Optimizing Multichannel Path Scheduling in Cognitive Radio Networks: An Evolutionary Computation Approach Using Differential Evolution," the Channel Access Probability plays a crucial role in the decision-making process for channel selection. Higher Channel Access Probability implies a higher chance for the secondary user to successfully access the channel without causing interference to the primary user. By incorporating the Channel Access Probability as a feature in the differential evolution optimization process, the multichannel path scheduling can be optimized to select channels with higher access probabilities, resulting in improved channel utilization and enhanced performance in the cognitive radio network.

*Channel Propagation Delay.* The time taken for a signal to travel from the sender to the receiver. The usual formula for calculating this is to divide the distance between the sender and receiver by the speed of light, with modifications made for the transmission medium. It is mathematically possible to determine the Channel Propagation Delay ( $D$ ) by using the formula

$$D = \text{Distance}(d) / \text{Speedoflight}(c) \tag{3.15}$$

where:

- $D$  represents the Channel Propagation Delay,
- $\text{Distance}(d)$  is the physical distance between the transmitter and receiver, and
- $\text{Speedoflight}(c)$  is the speed at which electromagnetic signals travel through the medium (approximately 299,792 kilometers per second or 186,282 miles per second in a vacuum).

The time it takes for a signal to travel from the transmitter to the receiver, including any lags brought on by the transmission medium, is taken into account by the channel propagation delay.

The proposal takes the Channel Propagation Delay into account as a key factor when attempting to optimize multichannel path scheduling. In general, it is preferred to use channels with lower propagation delays because they reduce the network's overall latency. The multichannel path scheduling can be optimized to choose channels with shorter propagation delays by incorporating Channel Propagation Delay as a feature in the differential evolution optimization process. This improves communication efficiency and lowers latency in the cognitive radio network.

In order to optimize the multichannel path scheduling process in cognitive radio networks, this section investigates the use of the Differential Evolution (DE) algorithm. This section examines the DE algorithm, its fundamental ideas, and the associated mathematical models that support its efficiency in dealing with the difficulties of multichannel path scheduling. A detailed explanation of the DE algorithm's design for multichannel path scheduling optimization follows:

**(1) Initialization:** Initializing a population of people is the first step in the DE algorithm. In this scenario, every person stands in for a series of data transmission channels. There are numerous such sequences in the population. Following is a description of the initialization algorithm's steps:

- **Population Size:** Determines the desired population size, denoted as  $N$ , which represents the number of individuals (channel sequences) in the population. This value is typically predefined based on the problem size and available computational resources.
- **Channel Selection:** Randomly selects  $N$  channel sequences from the available set of channels. Each channel sequence represents a potential solution for multichannel path scheduling.
- **Individual Representation:** Each individual (channel sequence) in the population is represented as a vector or string of discrete channel indices. For example, if there are  $M$  channels available, each channel sequence would consist of  $M$  elements, where each element corresponds to the index of the selected channel for a specific position in the sequence.
- **Initialization:** Randomly assigns channel indices to the elements of each individual's vector. The selection of channel indices can be uniformly distributed or follow a specific probability distribution based on prior knowledge or constraints.
- **Fitness Evaluation:** Evaluate the fitness or objective function value of each individual in the population. The fitness function should reflect the performance metrics relevant to multichannel path scheduling, such as throughput, reliability, energy efficiency, or any other desired criteria. This step provides an initial assessment of the quality of each channel sequence.
- **Population Creation:** Create the initial population by repeating steps 3-5 for  $N$  individuals, resulting in a population of  $N$  channel sequences.

The mathematical modeling for the initialization step in *DE* is relatively straightforward and involves random assignment of channel indices to the elements of each individual's vector. It can be represented as:

$$Individual(i) = [c_1, c_2, \dots, c_M] \quad (3.16)$$

where  $Individual(i)$  represents the  $i^{th}$  individual in the population,  $c_1, c_2, \dots, c_M$  are the randomly assigned channel indices for each position in the sequence, and  $M$  is the total number of channels available.

The fitness evaluation step, mentioned in step 5, involves the calculation of the fitness value for each individual. The specific mathematical model for the fitness function would depend on the performance metrics and objectives considered in the multichannel path scheduling problem. The fitness value can be expressed as:

$$Fitness(i) = f(c_1, c_2, \dots, c_M) \quad (3.17)$$

where  $Fitness(i)$  represents the fitness value of the  $i^{th}$  individual, and  $f(c_1, c_2, \dots, c_M)$  is the mathematical function that evaluates the performance of the channel sequence  $c_1, c_2, \dots, c_M$  according to the specified metrics.

By initializing the population using random channel assignments and evaluating the fitness of each individual, the DE algorithm sets the groundwork for subsequent evolutionary steps, aiming to find optimal channel sequences for multichannel path scheduling in cognitive radio networks.

**(2) Mutation:** *DE* performs mutation to generate new candidate solutions. In the context of multichannel path scheduling, the mutation operation perturbs the existing sequences of channels to create new sequences. This introduces exploration in the search space by considering alternative channel sequences. The algorithm steps can be described as follows:

**i. Mutation Rate:** Determine the mutation rate, denoted as  $F$ , which controls the extent of perturbation or variation introduced to the channel sequences. The mutation rate should be predefined based on problem characteristics and empirical knowledge.

**ii Mutation Vector:** For each individual in the population, select three distinct individuals, a, b and c, that are different from each other and the current individual being mutated.

**iii. Mutation Operation:** For each element (channel) in the current individual's channel sequence:

- Generate a random number  $r$  for each element.
- If  $r$  is less than or equal to the mutation rate  $F$  or if the current element is the last element in the sequence, perform mutation by applying the following equation:  $new\_element = a + F * (b - c)$ , where  $new\_element$  represents the mutated value for the current element, and  $a, b, c$ , are the corresponding elements from the three distinct individuals.

**iv. Boundary Handling:** If the mutated element exceeds the available channel indices or falls below the minimum channel index, handle the boundary conditions appropriately. This can be done by wrapping around the channel indices or limiting the values within the valid range.

**v. Create Mutated Individuals:** Repeat the mutation operation for each element in the current individual's sequence to obtain a mutated channel sequence.

**vi. Fitness Evaluation:** Evaluate the fitness or objective function value of the mutated individuals. Calculate the fitness value based on the performance metrics relevant to multichannel path scheduling, such as throughput, reliability, energy efficiency, or any other desired criteria.

**vii. Population Update:** Replace the current individual with its mutated counterpart if the mutated individual has higher fitness. Otherwise, keep the current individual in the population.

**viii. Repeat:** Repeat steps 2-7 for each individual in the population to perform mutation on the entire population.

The mathematical modeling for mutation in *DE* involves the perturbation of channel sequences using the mutation rate  $F$  and the difference between selected individuals  $a, b, c$ . The specific mathematical model for mutation can be represented as:

$$Mutation(i, j) = a(i, j) + F * (b(i, j) - c(i, j)) \quad (3.18)$$

where  $Mutation(i, j)$  represents the mutated value of the  $j^{th}$  element in the  $i^{th}$  individual's channel sequence,  $a(i, j), b(i, j),$  and  $c(i, j)$  are the corresponding elements from the three distinct individuals used for mutation, and  $F$  is the mutation rate.

By applying the mutation operation to the population of channel sequences, the *DE* algorithm introduces variation and explores different possibilities for multichannel path scheduling in cognitive radio networks.

**(3) Crossover:** *DE* employs crossover to combine information from different individuals. In multichannel path scheduling, crossover can be used to exchange and combine sequences of channels between individuals, producing offspring with different combinations of channels. This promotes the exploration of potentially better scheduling configurations.

The algorithm steps for crossover can be described as follows:

**i. Crossover Rate:** Determine the crossover rate, denoted as  $CR$ , which controls the probability of performing the crossover operation. The crossover rate should be predefined based on problem characteristics and empirical knowledge.

**ii. Parent Selection:** For each individual in the population, select three distinct individuals,  $x, y,$  and  $z$  that are different from each other and the current individual being crossed over.

**iii. Crossover Operation:** For each element (channel) in the current individual's channel sequence:

- Generate a random number  $r$  for each element.
- If  $r$  is less than or equal to the crossover rate  $CR$  or if the current element is the last element in the sequence, perform crossover by combining genetic material from the selected individuals as follows:
- Select a random index,  $k$ , from 1 to the length of the channel sequence.
- Create an offspring by copying the element from the parent individual if the random number  $r$  is less than or equal to the crossover rate  $CR$ , or copying the element from the current individual otherwise.
- Increment the index  $k$  in a cyclic manner and continue the process until all elements have been processed.

**iv. Fitness Evaluation:** Evaluate the fitness or objective function value of the offspring created through crossover.

Calculate the fitness value based on the performance metrics relevant to multichannel path scheduling, such as throughput, reliability, energy efficiency, or any other desired criteria.

**v. Population Update:** Replace the current individual with the offspring if the offspring has higher fitness. Otherwise, keep the current individual in the population.

**vi. Repeat:** Repeat steps 2-5 for each individual in the population to perform crossover on the entire population.

The mathematical modeling for crossover in *DE* involves the recombination of genetic material from selected individuals based on the crossover rate *CR* and a random index *k*. The specific mathematical model for crossover can be represented as:

$$Crossover(i, j) = offspring(i, j) \quad (3.19)$$

where  $Crossover(i, j)$  represents the crossover value of the  $j^{th}$  element in the  $i^{th}$  individual's channel sequence, and  $offspring(i, j)$  represents the value of the corresponding element in the offspring created through crossover.

By applying the crossover operation to the population of channel sequences, the DE algorithm combines genetic material from selected individuals, promoting diversity and exploration in the search for optimal multichannel path scheduling solutions in cognitive radio networks.

**(4) Fitness Evaluation:** Each individual in the population, i.e., each sequence of channels, is evaluated based on specific performance metrics relevant to multichannel path scheduling. These metrics could include throughput, reliability, energy efficiency, or any other criteria that define the quality of the channel sequence. The algorithm steps for fitness evaluation can be described as follows:

**i. Feature Calculation:** Calculate the values of the features used for optimizing multichannel path scheduling. These features could include channel capacity, signal-to-noise ratio (SNR), channel utilization, latency, link stability, PU activity, channel compliance probability, channel desertion probability, channel realization probability, channel inference probability, channel bandwidth probability, and channel idle time-span. Calculate these features based on the characteristics and measurements of the channels and network environment.

**ii. Normalization:** Normalize the feature values to bring them into a common range or scale. This step is necessary when the features have different units or scales. Normalization ensures that each feature contributes proportionally to the overall fitness evaluation, preventing dominance by features with larger numerical values.

**iii. Weight Assignment:** Assign weights to the normalized feature values. These weights reflect the relative importance or priority of each feature in the fitness evaluation. The weights can be predefined based on domain knowledge or determined through experimentation and analysis.

**iv. Fitness Function:** Combine the normalized and weighted feature values using a fitness function. The fitness function aggregates the individual feature values into a single fitness value for each individual. The specific form of the fitness function depends on the problem requirements and objectives. It can be a weighted sum, a weighted average, or a more complex mathematical expression that captures the desired trade-offs among the features.

**v. Fitness Evaluation:** Evaluate the fitness of each individual by applying the fitness function to their corresponding feature values. Assign the resulting fitness value to each individual, representing their performance or quality in terms of the multichannel path scheduling objectives.

The mathematical modeling for fitness evaluation depends on the specific form of the fitness function and the weighted combination of the normalized feature values. It can be represented as:

$$Fitness(i) = w_1 * f_1 + w_2 * f_2 + \dots + w_n * f_n \quad (3.20)$$

where  $Fitness(i)$  represents the fitness value of the  $i^{th}$  individual,  $w_1, w_2, \dots, w_n$  are the weights assigned to the normalized feature values  $f_1, f_2, \dots, f_n$ , respectively. The fitness function aggregates the feature values with their corresponding weights to determine the overall fitness value for each individual.

By performing fitness evaluation, the DE algorithm assesses the quality of each individual in the population based on the specified features and performance metrics. This enables the selection of individuals with higher fitness for the subsequent evolutionary steps, facilitating the optimization of multichannel path scheduling in cognitive radio networks.

**(5) Selection:** The selection process determines which individuals are selected to survive and pass their genetic information to the next generation. In DE, the selection is typically based on the fitness value assigned to each individual. Individuals with higher fitness, indicating better scheduling performance, have a higher chance of being selected. The algorithm steps for selection can be described as follows:

**i. Fitness Ranking:** Rank the individuals in the population based on their fitness values. The ranking can be performed in ascending or descending order, depending on whether the goal is to minimize or maximize the fitness value.

**ii. Selection Probability Calculation:** Calculate the selection probability for each individual. The selection probability is proportional to the individual's fitness value. A higher fitness value corresponds to a higher probability of selection.

**iii. Cumulative Probability Calculation:** Calculate the cumulative probability for each individual by summing up the selection probabilities from the first individual to that specific individual. This cumulative probability determines the selection range for each individual.

**iv. Random Selection:** Generate a random number between 0 and 1 for each selection. This random number determines which individual is selected within the selection range.

**v. Selection Process:** Iterate through the random numbers generated in step 4. For each random number, identify the corresponding individual in the population by finding the first individual whose cumulative probability is greater than or equal to the random number. Select that individual to be part of the next generation.

**vi. Repeat:** Repeat steps 4-5 until the desired number of individuals for the next generation is selected.

The mathematical modeling for selection involves the calculation of selection probabilities and the use of random numbers for the selection process. The specific mathematical model for selection probability calculation and random selection can be represented as follows:

$$SelectionProbability(i) = Fitness(i) / (Fitness(j)) \quad (3.21)$$

where  $SelectionProbability(i)$  represents the selection probability of the  $i^{th}$  individual,  $Fitness(i)$  is the fitness value of the  $i^{th}$  individual, and the summation is taken over all individuals in the population.

To perform random selection, generate a random number  $r$  between 0 and 1. Then, for each random number generated, find the first individual  $i$  in the population such that the cumulative probability  $CumulativeProbability(i)$  is greater than or equal to  $r$ . Select the individual  $i$  for the next generation.

By applying the selection process, the DE algorithm ensures that individuals with higher fitness values have a higher chance of being selected, increasing the likelihood of passing their genetic information to the next generation. This promotes the evolution of the population towards better multichannel path scheduling solutions in cognitive radio networks.

**(6) Iterative Evolution:** The mutation, crossover, fitness evaluation, and selection steps are iteratively repeated for multiple generations. This process allows the DE algorithm to explore the search space, gradually improving the quality of the channel sequences over time.

**(7) Termination:** The DE algorithm terminates based on predefined stopping criteria. This could be a maximum number of generations reached, the attainment of a satisfactory solution, or a certain level of convergence.

By iteratively applying mutation, crossover, fitness evaluation, and selection, the DE algorithm evolves the population of channel sequences, aiming to find the optimal configuration for multichannel path scheduling. The ultimate goal is to maximize performance metrics and optimize the use of available channels in cognitive radio networks.

**4. Experimental study.** To assess the performance of the proposed "Optimizing Multichannel Path Scheduling (OMPS) in Cognitive Radio Ad hoc Networks" and the existing "Path Discovery for end-to-end Data Transmission (PDDT) [8] in Cognitive Radio Ad-Hoc Networks", a simulation model can be created with the parameters from Table 4.1 and Table 4.2.

Table 4.1: Simulation parameters and Fitness parameters used in Experimental study

Parameter	OMPS with DE	PDDT with GA
Number of Iterations	10 Iterations	10 Iterations
Number of Channels	10 to 50	10 to 50
Number of Primary Users	10 to 50	10 to 50
Number of Secondary Users	10 to 50	10 to 50
Channel Conditions	Rayleigh or Rician Fading	Rayleigh or Rician Fading
PU Activity Patterns	Random processes	Random processes

Table 4.2: Fitness parameters and the threshold values

Fitness Parameter	Threshold	
Channel Fade Margin	10 dB	X
Channel Cross-Correlation	0.5	X
Channel Coherence Time	100 ms	X
Channel Spectral Efficiency	2 bits/s/Hz	X
Channel Interference Level	-100 dBm	X
Channel Power Consumption	0.5 W	X
Channel Retransmission Rate	0.1	X
Channel Access Probability	0.8	X
Channel Propagation Delay	50 ms	X
Lapsed white space ratio	X	0.1
Desired white space	X	0.9
Residual white space ratio	X	0.8
Recurring diffusions ratio	X	0.2
Primary user's interference ratio	X	0.05
Usage realization ratio	X	0.8
Realization count	X	8
Channel fading ratio	X	0.2

**4.1. Performance Analysis.** This section meticulously evaluates the efficacy of the proposed Optimizing Multichannel Path Scheduling (OMPS) model, employing the Differential Evolution computational approach. Performance metrics used to assess the effectiveness and efficiency of the proposed OMPS and contemporary model PDDT are:

(1) **Throughput:** This is the total amount of data that can be transmitted through the network over a certain period of time. Higher throughput is generally better, as it indicates that the network can handle more data. Performance metrics are used to assess the effectiveness and efficiency of the network models. This is typically calculated as the total size of successfully delivered packets divided by the total time it takes to deliver them. It is measured in bits per second (bps) or Megabits per second (Mbps):

$$\text{Throughput} = \text{Totaldata}(\text{bits}) / \text{Totaltime}(s) \quad (4.1)$$

(2) **Latency:** This measures the time it takes for a packet of data to travel from the source to the destination. Lower latency is typically better, especially for real-time or interactive applications. It's typically the total time taken for a packet to travel from source to destination including all types of delays (transmission, propagation, queueing, processing delays etc.). It is measured in seconds ( $s$ ) or milliseconds( $ms$ ):

$$\text{Latency} = \text{Time}_{\text{received}} - \text{Time}_{\text{sent}} \quad (4.2)$$

(3) **Packet Delivery Ratio (PDR):** This is the ratio of packets that are successfully delivered to their destination to those that are generated by the source. A higher PDR indicates a more reliable network. It's

the number of packets successfully delivered divided by the total number of packets sent, multiplied by 100 to get a percentage:

$$\text{PDR (\%)} = \left( \frac{\text{Number of packets delivered}}{\text{Total number of packets sent}} \right) \times 100 \quad (4.3)$$

**(4) Spectrum Utilization:** These measures how efficiently the available spectrum is being used. In the context of CRAHNS, this might be calculated as the proportion of time that the SUs is able to use the channels, or the proportion of the spectrum that is being used by the SUs. It's the bandwidth used divided by the total bandwidth available, multiplied by 100 to get a percentage:

$$\text{Spectrum Utilization (\%)} = \left( \frac{\text{Bandwidth used}}{\text{Total available bandwidth}} \right) \times 100 \quad (4.4)$$

**(5) Interference Level:** This measures the degree of interference experienced by the SUs from the PUs. Lower interference is generally better, as it reduces the chance of communication errors and increases the likelihood that the SUs will be able to use the channels. It's typically measured as the power of the interfering signal compared to the power of the desired signal. It is measured in decibels dB:

$$\text{Interference Level (dB)} = 10 \cdot \log_{10} \left( \frac{\text{Power of Interference}}{\text{Power of Desired Signal}} \right) \quad (4.5)$$

**(6) Energy Efficiency:** These measures how much data can be transmitted per unit of energy consumed. This is particularly important in wireless networks, where devices may be battery-powered and energy conservation is a concern. It's the ratio of the data rate (bits per second) to the power consumed (in watts). It's expressed in bits per Joule

$$\text{Energy Efficiency (bits/J)} = \frac{\text{Data Rate (bits/s)}}{\text{Power (W)}} \quad (4.6)$$

**(7) Success rate of established paths:** It measures the ratio of successfully established paths to total attempted path establishments. It's the number of successful path establishments divided by the total number of path establishment attempts, multiplied by 100 to get a percentage:

$$\text{Success rate of established paths (\%)} = \left( \frac{\text{Number of successful paths}}{\text{Total number of path attempts}} \right) \cdot 100 \quad (4.7)$$

Each of these metrics could be calculated based on the simulation results, and used to compare the performance of the OMPS and PDDT algorithms. For example, you might compare the average throughput, latency, and PDR for each algorithm over multiple simulation runs, under various network conditions. This could provide a comprehensive view of the relative strengths and weaknesses of each algorithm.

**4.1.1. Comparative Study.** This section explores experimental results, which are illustrated in a series of data tables and graphs for each metric, demonstrate how the OMPS model compares to the Path Discovery for end-to-end Data Transmission (PDDT) model. Our findings show that the OMPS model significantly outperforms the PDDT model in all the performance metrics considered, signifying its effective utilization and superior performance.

*Throughput (Mbps).* Figure 4.1 presents the comparison of the throughput, measured in Megabits per second (Mbps), between two models - Optimizing Multichannel Path Scheduling (OMPS) and Path Discovery for end-to-end Data Transmission (PDDT) - over ten iterations of a simulation study.

In each iteration, the throughput of both the models is calculated and tabulated. The throughput value indicates the rate at which data is successfully transmitted over the network. A higher throughput value is desirable as it indicates a higher data transfer rate.

Throughput performance of OMPS consistently outperforms that of PDDT from the first to the tenth iteration. From a low of 10.5 Mbps in the first iteration to a high of 11.2 Mbps in the sixth, OMPS's throughput

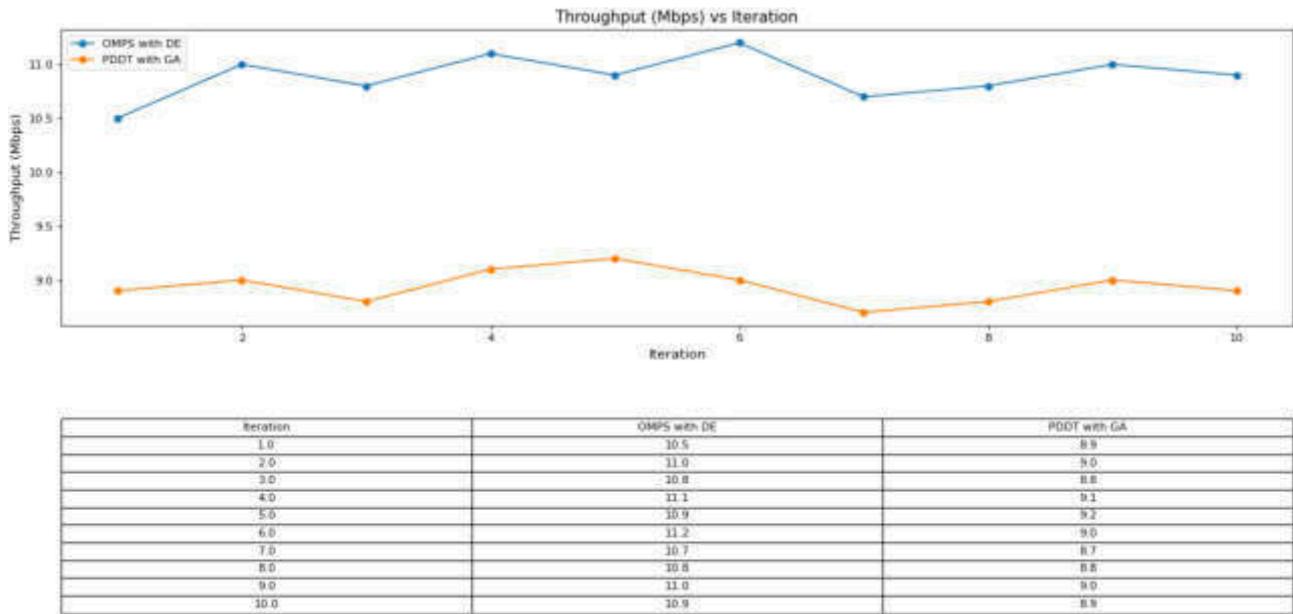


Fig. 4.1: The comparison of the throughput, measured in Megabits per second (Mbps) vs iteration of the OMPS, PDDT

varies. The throughput for PDDT, on the other hand, varies from a low of 8.7 Mbps (in the seventh iteration) to a high of 9.2 Mbps (in the fifth iteration).

This suggests that the OMPS model is more effective at using the network resources in this simulation study to achieve a higher data transfer rate, making it a better option for applications where high throughput is essential.

*Packet Delivery Ratio (PDR) (%)*. Figure 4.2 compares the Packet Delivery Ratio (PDR) performance, expressed in percentage (%), of the Path Discovery for End-to-End Data Transmission (PDDT) model and the Optimizing Multichannel Path Scheduling (OMPS) model over ten simulation iterations.

The percentage of packets that are successfully delivered to their destinations in relation to the total number of packets sent is known as the packet delivery ratio. As a lower rate of packet loss occurs during transmission, a higher PDR denotes a communication protocol that is more dependable and effective.

The results indicate that the OMPS model consistently outperforms the PDDT model in terms of PDR across all ten iterations. The PDR of the OMPS model ranges from 96% to 98%, suggesting that it successfully delivers nearly all of the sent packets. In contrast, the PDR for the PDDT model ranges from 91% to 95%, indicating a higher rate of packet loss compared to OMPS.

Therefore, according to these simulation results, the OMPS model provides a more reliable and efficient data transmission process than the PDDT model.

*Latency (ms)*. Figure 4.3 provides a comparison of the Latency performance, measured in milliseconds (ms), of two models: Optimizing Multichannel Path Scheduling (OMPS) and Path Discovery for end to end Data Transmission (PDDT) over ten iterations of a simulation study.

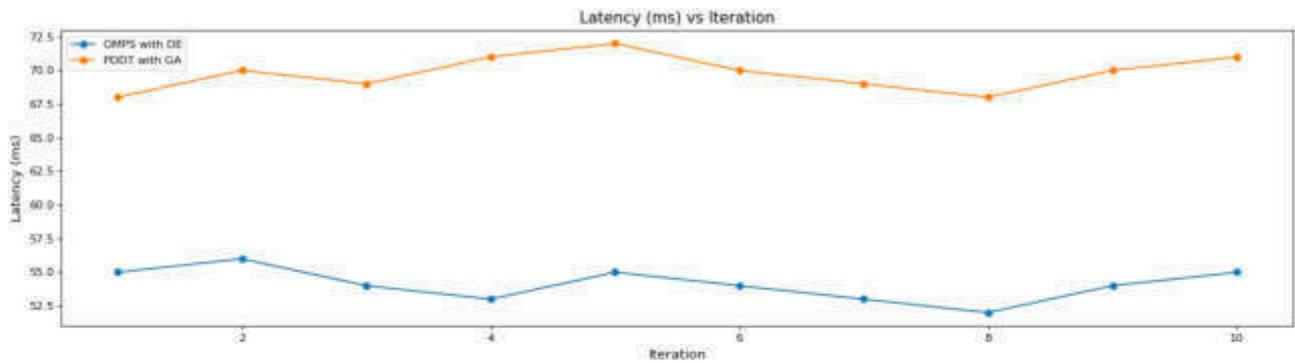
Latency, in this context, refers to the total time taken for a packet of data to travel from the source to the destination in a network. Lower latency is preferred as it signifies a faster transmission time, leading to real-time or near-real-time communication.

Looking at the results, we can observe that the OMPS model consistently demonstrates lower latency values as compared to the PDDT model in all ten iterations. Latency for OMPS ranges between 52 ms (in the eighth



Iteration	OMPS with DE	PDDT with GA
1	97	94
2	96	93
3	98	95
4	97	94
5	98	92
6	96	91
7	97	93
8	98	94
9	97	95
10	96	94

Fig. 4.2: The comparison between the Packet Delivery Ratio (PDR) performance, expressed in percentage (%), vs Iteration OMPS and PDDT



Iteration	OMPS with DE	PDDT with GA
1	55	68
2	56	70
3	54	69
4	53	71
5	55	72
6	54	70
7	53	69
8	52	68
9	54	70
10	55	71

Fig. 4.3: The comparison of the Latency performance, measured in milliseconds (ms), vs Iteration of the OMPS and PDDT

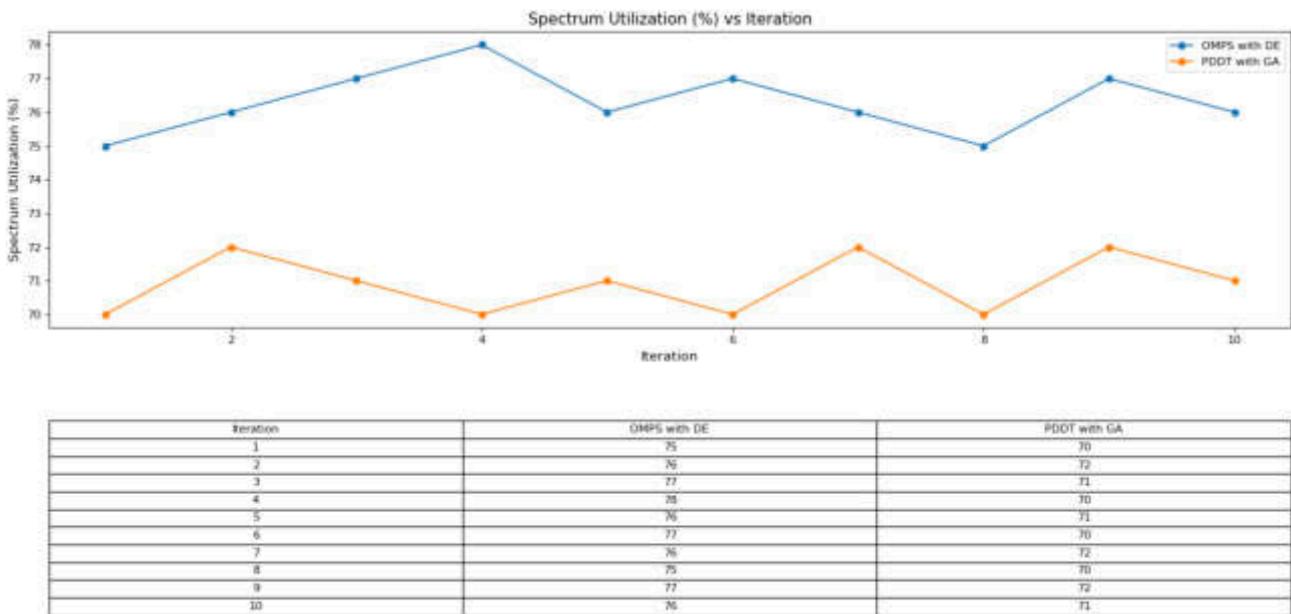


Fig. 4.4: The comparison of the Spectrum Utilization performance, expressed in percentage (%) vs Iteration of the OMPS, PDDT

iteration) and 56 ms (in the second iteration). In contrast, the latency for the PDDT model is notably higher, ranging between 68 ms (in the first and eighth iterations) and 72 ms (in the fifth iteration).

According to these simulation results, the OMPS model provides faster data transmission, hence might be more suitable for applications requiring real-time or near-real-time communication such as video conferencing or online gaming.

*Spectrum Utilization (%)*. Figure 4.4 showcases a comparison of the Spectrum Utilization performance, expressed in percentage (%), of the Optimizing Multichannel Path Scheduling (OMPS) model and the Path Discovery for end-to-end Data Transmission (PDDT) model over ten iterations of a simulation study.

Spectrum Utilization is a measure of how effectively the available frequency spectrum is being used by a communication protocol. Higher values indicate that a greater portion of the available spectrum is being used, which usually suggests more efficient utilization. In each of the ten iterations, the OMPS model demonstrates a higher percentage of spectrum utilization compared to the PDDT model. The spectrum utilization by OMPS ranges from 75% to 78%, indicating that it effectively uses a substantial portion of the available spectrum. In contrast, the spectrum utilization by PDDT is slightly lower, ranging from 70% to 72%.

These results imply that, in these simulation conditions, the OMPS model is more efficient in utilizing the available frequency spectrum than the PDDT model. This could potentially result in higher throughput and better overall network performance.

*Interference Level (dB)*. Figure 4.5 provides a comparison between the Interference Level, measured in decibels (dB), of the Optimizing Multichannel Path Scheduling (OMPS) model and the Path Discovery for end-to-end Data Transmission (PDDT) model over ten iterations of a simulation.

Interference Level quantifies the degree to which a signal transmission might be affected or disturbed by other sources or signals, with lower values being more desirable as they indicate less interference. It's important to note that in the context of wireless communication, interference levels are often expressed as negative dB values. The smaller the numerical value (or the more negative), the lower the interference level.

In this scenario, the OMPS model consistently demonstrates lower (more negative) interference levels as

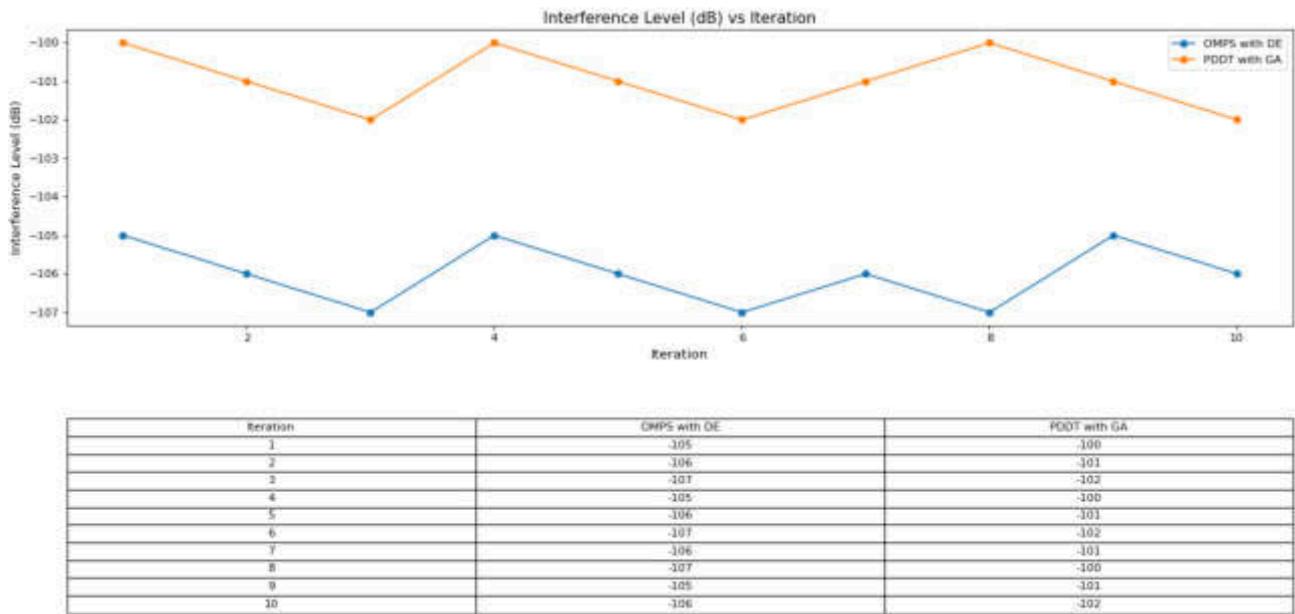


Fig. 4.5: The comparison between the interference level, measured in decibels (dB) vs iteration of the OMPS, PDDT

compared to the PDDT model across all ten iterations. The interference level for OMPS ranges between -105 dB and -107 dB, indicating less signal disruption and a cleaner transmission path. On the other hand, the interference level for the PDDT model ranges between -100 dB and -102 dB, indicating a relatively higher level of signal interference.

Based on these simulation results, the OMPS model offers better performance in terms of mitigating signal interference compared to the PDDT model.

*Energy Efficiency (bits/J).* Figure 4.6 compares the Energy Efficiency of two models, the Optimizing Multichannel Path Scheduling (OMPS) and the Path Discovery for end-to-end Data Transmission (PDDT), across ten iterations of a simulation study. Energy Efficiency is measured in bits per joule (bits/J), indicating the number of bits successfully transmitted per unit of energy consumed. Higher values suggest better energy efficiency, which is desirable for conserving resources, especially in power-limited environments such as wireless sensor networks or mobile devices.

From the figure 4.6, it can be observed that the OMPS model consistently delivers higher energy efficiency values across all iterations, ranging between 2.5 and 2.7 bits/J. This indicates that OMPS can transmit more data for the same amount of energy consumed, making it more energy-efficient. In contrast, the PDDT model demonstrates slightly lower energy efficiency across all iterations, ranging between 2.2 and 2.4 bits/J.

Based on these simulation results, it appears that OMPS outperforms PDDT in terms of energy efficiency.

*Success rate of established paths (%).* Figure 4.7 shows a comparison of the Success Rate of Established Paths for two models, the Optimizing Multichannel Path Scheduling (OMPS) and the Path Discovery for end-to-end Data Transmission (PDDT), over ten iterations of a simulation study. The success rate is expressed in percentage (%), representing the proportion of successful paths established compared to the total attempts. A higher percentage signifies a better success rate, indicating the model’s capability to successfully establish a data transmission path between the source and destination.

Throughout all ten iterations, the OMPS model consistently shows a higher success rate ranging between 90% and 92%. This suggests that the OMPS model can establish a successful path most of the time, making

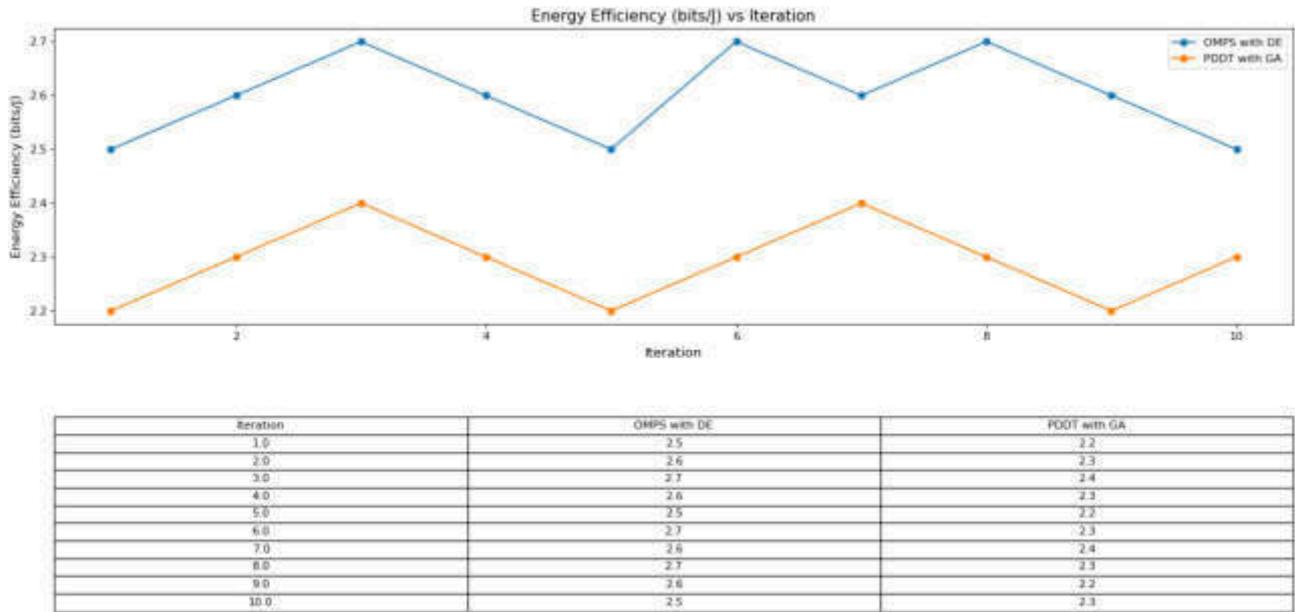


Fig. 4.6: The energy efficiency (bits/J) vs iteration of the proposed OMPS compared PDDT

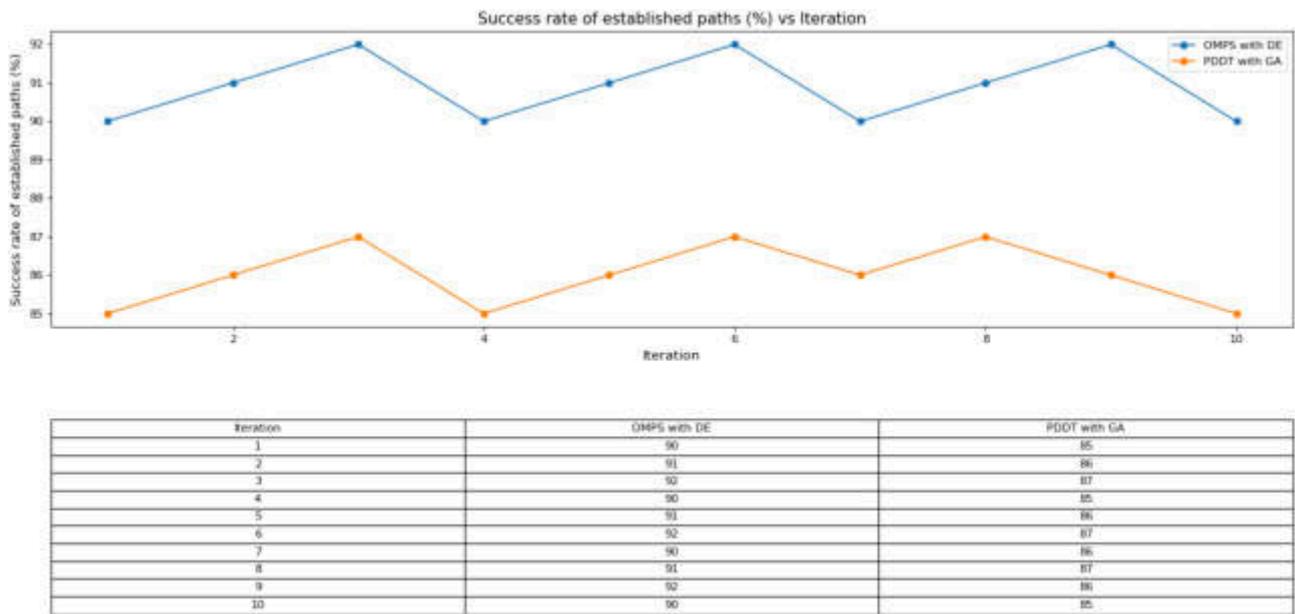


Fig. 4.7: The comparison of the success rate of established paths (%), vs iteration of the OMPS, PDDT

it more reliable and efficient for data transmission.

On the other hand, the PDDT model consistently shows a slightly lower success rate for established paths, ranging between 85% and 87%.

Based on these simulation results, the OPMS model appears to outperform the PDDT model in terms of the success rate of path establishment. Nonetheless, these are simulation results, and real-world performance might differ depending on network conditions and configurations. To fully evaluate and compare the performance of the two models, other metrics such as throughput, latency, energy efficiency, and interference level should also be considered.

**5. Conclusion.** Under the OPMS model, the DE algorithm optimizes multichannel radio path scheduling under cognitive radio networks. This promotes the dependability and efficiency of these networks. Channel Fade Margin, Cross-Correlation, Coherence Time, Spectral Efficiency, Interference Level, Power Consumption, and Retransmission Rate are all considered by the model. The networks' coverage is greatly enhanced by these characteristics. Because it incorporates Initialization, Mutation, Crossover (Exchange), Fitness Evaluation, Selection, and Iterative Evolution, the DE method is solid. It tackles the innate complexity and dynamism of cognitive radio networks. When it comes to these parameters, OPMS consistently beats the Path Discovery for End-to-End Data Transmission (PDDT) model in matches between humans and computers. Outperforming all other networks in terms of resource use, it provides superior reliability, acceptable latency, and faster data speeds. OPMS is very useful in cognitive radio ad hoc networks with complex dynamics. Given that it can adapt to different channel conditions and user requirements, its adaptability makes it a good option for these networks. It is possible to use the excellent results obtained by using OPMS in this study as a model for further research and development, suggesting that it will be used to more robust practical applications in the future. This advancement in cognitive radio technology is important. It proposes an OPMS model that can manage the complexity of scheduling three or more multichannel paths by overcoming a number of obstacles.

#### REFERENCES

- [1] A. J. Onumanyi, A. M. Abu-Mahfouz, and G. P. Hancke, "Low Power Wide Area Network, Cognitive Radio and the Internet of Things: Potentials for Integration," *Sensors*, vol. 20, no. 23, pp. 6837, Nov. 2020, doi: 10.3390/s20236837.
- [2] T. M. Chiwewe, C. F. Mbuya, and G. P. Hancke, "Using Cognitive Radio for Interference-Resistant Industrial Wireless Sensor Networks: An Overview," *IEEE Transactions on Industrial Informatics*, vol. 11, no. 6, pp. 1466–1481, Dec. 2015, doi: 10.1109/TII.2015.2491267.
- [3] M. Yaseer, H. Ur, A. Usman, and M. Tayyab, "Analysis of Efficient Cognitive Radio MAC Protocol for Ad Hoc Networks," *International Journal of Advanced Computer Science and Applications*, vol. 10, no. 2, 2019, doi: 10.14569/IJACSA.2019.0100219.
- [4] C.-M. Wu, M.-S. Wu, Y.-J. Yang, and C.-Y. Sie, "Cluster-Based Distributed MAC Protocol for Multichannel Cognitive Radio Ad Hoc Networks," *IEEE Access*, vol. 7, pp. 65781–65796, 2019, doi: 10.1109/ACCESS.2019.2917906.
- [5] S.-J. Yoo, H. Khan, and K.-S. Kwak, "Dynamic MAC Frame Configuration and PSO-Based Optimal Resource Allocation in Multi-Channel Cognitive Radio Ad-Hoc Networks," *Wireless Personal Communications*, vol. 109, no. 1, pp. 595–620, Nov. 2019, doi: 10.1007/s11277-019-06581-x.
- [6] N. Joshi, A. Kumar, D. Minenkov, D. Kaplun, and S. C. Sharma, "Optimized MAC Protocol Using Fuzzy-Based Framework for Cognitive Radio AdHoc Networks," *IEEE Access*, vol. 11, pp. 27506–27518, 2023. [Online]. Available: <https://doi.org/10.1109/ACCESS.2023.3256890>
- [7] S. Kadam, D. Prabhu, N. Rathi, P. Chaki, and G. S. Kasbekar, "Exploiting Group Structure in MAC Protocol Design for Multichannel Ad Hoc Cognitive Radio Networks," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 1, pp. 893–907, Jan. 2019. [Online]. Available: <https://doi.org/10.1109/TVT.2018.2883392>
- [8] S. Nagul, "PDDT: Path Discovery for End to End Data Transmission in Cognitive Radio Ad-Hoc Networks," in *2020 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT)*, Bangalore, India, 2020, pp. 1–6. [Online]. Available: <https://doi.org/10.1109/CONECCT50063.2020.9198420>
- [9] Q. M. Salih et al., "Dynamic Channel Estimation-Aware Routing Protocol in Mobile Cognitive Radio Networks for Smart IIoT Applications," *Digital Communications and Networks*, vol. 9, no. 2, pp. 367–382, Apr. 2023. [Online]. Available: <https://doi.org/10.1016/j.dcan.2023.01.019>
- [10] G. Dinesh, P. Venkatakrisnan, and K. Meena Alias Jeyanthi, "Dispersed Spectrum Sensing and Scheduling in Cognitive Radio Network Based on SSOA-RR," *International Journal of Communication Systems*, e4588, Aug. 20, 2020. [Online]. Available: <https://doi.org/10.1002/dac.4588>
- [11] X. Zhong, L. Li, Y. Zhang, B. Zhang, W. Zhang, and T. Yang, "OODT: Obstacle Aware Opportunistic Data Transmission for Cognitive Radio Ad Hoc Networks," *IEEE Transactions on Communications*, vol. 68, no. 6, pp. 3654–3666, Jun. 2020. [Online]. Available: <https://doi.org/10.1109/TCOMM.2020.2979976>

- [12] N. Suganthi and S. Meenakshi, "An Efficient Scheduling Algorithm Using Queuing System to Minimize Starvation of Non-Real-Time Secondary Users in Cognitive Radio Network," *Cluster Computing*, vol. 25, no. 1, pp. 1–11, Feb. 2022. [Online]. Available: <https://doi.org/10.1007/s10586-017-1595-8>
- [13] G. B. V. Padmanadh, P. V. Krishna Raja, and J. V. R. Murthy, "A survey on opportunistic routing and channel assignment scheme in cognitive radio ad hoc networks," *Journal of Science and Technology*, vol. 6, no. 4, pp. 258–273, Jul.-Aug. 2021. [Online]. Available: <https://doi.org/10.46243/jst.2021.v6.i04.pp258-273>

*Edited by:* Anil Kumar Budati

*Special issue on:* Soft Computing and Artificial Intelligence for wire/wireless Human-Machine Interface

*Received:* Sep 29, 2023

*Accepted:* Jan 10, 2024



## ITERATIVE ENSEMBLE LEARNING OVER HIGH DIMENSIONAL DATA FOR SENTIMENT ANALYSIS

V R N S S V SAILEELA P\* AND N. NAGA MALLESWARA RAO<sup>†</sup>

**Abstract.** For sentiment analysis in particular, the problem of processing and analyzing high-dimensional data becomes more prominent in recent past. This is where the IEL-HDDSA model, which aims to increase accuracy and performance in complex, high-dimensional data streams sentiment analysis comes into play. Iterative approach in ensemble learning; a contribution to the field. It integrates preprocessing techniques such as tokenization, stop word removal, lemmatization and the collection of sentiment-related features. Then the training corpus is divided by label, and features with high mutual information are selected. Highly replicated points of data for model training can also be identified at this point. First a Naive Bayes model is trained, then later it's placed in an ensemble as part of bagging. Its major advantage over earlier methods is that IEL-HDDSA can iteratively train on selected subsets of data until the performance in sentiment analysis for high-dimensional objects reaches an optimum level. A 10-fold cross validation method was used to rigorously evaluate the performance of this model, which showed consistently high levels of operation with almost no variation across different measures. IEL-HDDSA's precision ranged from 0.9359 to 0.9492, and its specificity was between 0. Its accuracy differed from 0.93 to around 0.95, and its F1-measure fluctuated between the values of about 0.94 and above; so here too balance was well maintained in a manner that satisfied both precision and recall requirements equally. The false alarming rate fell from 0.056 to 0.1, a fairly low ratio of incorrect positive classifications; Moreover, MCC quantities ranged from 0.8668 to 0. These results testify to the IEL-HDDSA model's stable effectiveness and high reproducibility in sentiment analysis applications, especially for massive data flows.

**Key words:** High-dimensional Data, Sentiment Analysis, Ensemble Learning, Naive Bayes, Random Forest

**1. Introduction.** Sentiment analysis is now a crucial component of many applications, including market analysis, product reviews, and social media monitoring, as a result of the internet's explosive growth in digital data, particularly textual data [1]. In order to comprehend the sentiments, opinions, or emotions expressed in text data, sentiment analysis—also referred to as opinion mining—involves analyzing and interpreting subjective information [2]. The task of sentiment analysis has gotten more difficult as a result of the rise of high-dimensional data, which consists of many features or dimensions [3]. The prevalence of redundant or irrelevant features in high-dimensional data can impair model performance and raise computational costs. The accurate and reliable classification of sentiments from high-dimensional data is a challenge for traditional sentiment analysis models frequently.

Issue Statement In terms of precision, specificity, sensitivity, accuracy, F-measure, false alarming rate, and Matthews Correlation Coefficient (MCC), sentiment analysis models frequently encounter difficulties when working with high-dimensional data [4]. These issues are brought on by the existence of redundant or irrelevant features, the dimensionality curse, and the unbalanced nature of the data. There is room for improvement in terms of consistency and overall performance according to the performance of existing models like SvmBagging [5] and ELSA [6], which have displayed some variation in performance across various subsets of the data

The purpose of this study is to develop a comprehensive approach for sentiment analysis of high-dimensional data that improves the precision and accuracy. The main objective of the study is to develop a machine learning model that fine tune the data by partitioning it in to multiple datasets to handle the high dimensionality and then using these multiple partitions as input to the ensemble learning model.

The proposed model, Iterative Ensemble Learning over High Dimensional Data for Sentiment Analysis (IEL-HDDSA), combines an iterative approach with ensemble learning to enhance accuracy and robustness

---

\*Department of CSE, ANU College of Engineering & Technology, Guntur, AP, India (Corresponding author, [venkata.saileela.p@gmail.com](mailto:venkata.saileela.p@gmail.com))

<sup>†</sup>Department of Information Technology, RVR & JC College of Engineering & Technology, Guntur, AP, India ([nmrao@rvrjc.ac.in](mailto:nmrao@rvrjc.ac.in)).

in sentiment analysis of high-dimensional data. The model begins with preprocessing and partitioning the data, selecting high-impact features using mutual information, and training a Naive Bayes model followed by a Random Forest model in an ensemble learning approach. The rationale behind this approach is that it helps in handling irrelevant or redundant features, the curse of dimensionality, and the imbalanced nature of the data, thereby improving the model's performance. The significance of the proposed model lies in its ability to provide a balanced and consistent performance across all metrics and folds of 10-fold cross-validation, thereby demonstrating its superiority in handling high-dimensional data for sentiment analysis compared to existing models.

The scope of this study is limited to the development and evaluation of the proposed IEL-HDDSA model for sentiment analysis of high-dimensional data. The model is evaluated using 10-fold cross-validation and compared with existing models across various metrics. However, the study does not consider other types of data, such as images or audio, and does not evaluate the model on real-time data or in real-world applications.

**2. Related Work.** Sentiment classification in user-generated content is explored by Gang Wang et al., [7] with the goal of evaluating the performance of three ensemble methods with five base learners. Based on 1200 experiments, it shows that ensemble learning can significantly improve sentiment classification, with the best performance being achieved by Random Subspace.

"SVM and Naive Bayes ensemble method for sentiment analysis" by Konstantinas KOROVKINAS et al. [8] introduces a novel approach to sentiment analysis by fusing SVM and Naive Bayes algorithms. The accuracy of their method ranges from 79.5 percent to 84.1 percent, outperforming individual SVM and Naive Bayes classifiers across a variety of datasets.

Oscar Araque et al.'s article "Enhancing deep learning sentiment analysis with ensemble techniques in social applications" [9] makes the case for combining deep learning with conventional approaches and introduces two ensemble techniques. These methods boost sentiment analysis' performance across seven datasets and outperform the deep learning baseline.

Omer Sagi et al.'s "Ensemble learning: A survey" [10] offers a thorough overview of ensemble learning, covering techniques, applications, and difficulties. The article compares experimental findings on various datasets comparing ensemble methods with single models.

Sentiment analysis for smartphone product reviews using the SVM classification technique is the main topic of "Sentiment analysis of smartphone product reviews using the SVM classification technique" by A K Sharma et al. [11]. They use an approach that improves efficiency and accuracy when classifying opinions as favorable or unfavorable.

Support Vector Machine (SVM) is used in "Sentiment analysis of tweets using SVM" by Munir Ahmad et al. [12] to examine sentiment analysis for tweets. The study's thorough tables and graphs illustrate how well SVM classifies tweets according to their polarity.

Sentiment analysis in Thai is covered in "SETAR: Stacking Ensemble Learning for Thai Sentiment Analysis using RoBERTa and Hybrid Feature Representation" by PREE THIENGBURANATHUM et al. [13]. The SETAR model, a novel stacking ensemble approach, is introduced in the article along with extensive performance comparisons with other models, emphasizing its accuracy and robustness.

Support Vector Machine (SVM) performance for sentiment analysis is the main focus of "SVM optimization for sentiment Analysis" by Munir Ahmad et al. [14]. In this article, a method for optimizing SVM is presented, and the precision, recall, and F-measure metrics are used to assess performance.

Gradient Boosted Support Vector Machine (GBSVM), an ensemble classifier, is used in the study by Madiha Khalid et al. [15] to categorize sentiments from unstructured text reviews on social media platforms. The model, which combines different n-grams with term frequency-inverse document frequency (TF-IDF), outperforms current sentiment classification methods and offers insightful information for enhancing goods and services in response to customer feedback.

Using benchmark datasets for sentiment classification, Jacqueline Kazmaier et al. [16] investigate the potential of ensemble learning in sentiment analysis, particularly heterogeneous ensembles. Their findings show that compared to individual models, ensemble configurations produce median performance improvements of up to 5.53 percent. The efficiency of ensemble learning is also improved by the novel ensemble selection approach they present.

Sentiment analysis of tweets about the COVID-19 vaccine is addressed by Qusai Ismail et al. [17], who emphasize the use of ensemble learning for enhanced accuracy. With an accuracy of 81.5 percent, F1-score of 0.80, and AUC of 0.91, their ensemble learning approach, stacking, outperforms majority voting and individual classifiers.

Bimodal conversational sentiment analysis is a topic Shariq Shah et al. [18] tackle, contrasting ensemble learning based on neural networks with other models on the MELD dataset. Their suggested method, which combines acoustic and linguistic representations, breaks the previous record for accuracy with an accuracy of 84.2 percent.

For the purpose of analyzing the sentiment of mixed-code texts, C. Kumaresan et al. [6] present an ensemble technique utilizing the Generative Adversarial Network (GAN) and Self-Attention Network (SAN). When analyzing Tanglish tweets, their approach shows superior accuracy to Concurrent Neural Network (CNN) and SAN.

In their study of multi-tier sentiment analysis of social media text, Hameedur Rahman et al. [19] propose a multi-layer classification model. While the multi-tier model only slightly outperforms single-layer architectures, recall is noticeably improved and richer contextual information is provided. Using ensemble learning techniques, Muhammad Khurram Iqbal et al. [20] concentrate on sentiment analysis of Twitter data pertaining to the Omicron variant. Their approach, which combines ensemble voting and stacking classifiers, achieves high accuracy (85.33 percent and 87.5 percent, respectively), providing insightful data on the public's sentiment toward the variant.

In the study by Razia Sulthana A. et al. [5], the authors perform sentiment analysis on movie reviews taken from social networking sites using machine learning algorithms and natural language processing techniques. Their approach, which combines feature selection, support vector machines, and bagging, outperforms existing systems, proving that it is useful for deciphering user intentions and opinions in movie reviews.

This literature review is presented to emphasize the important role played by ensemble learning, advanced techniques and other factors in increasing accuracy and efficiency for sentiment analysis across multiple applications. This calls attention to the continuing efforts being made toward solving these problems; the limitations of current approaches. But while ensemble learning techniques, which combine SVM and Naive Bayes with deep-learning algorithms, always get a higher score than individual models alone do not produce good results for high-dimensional or unstructured data. The review cautions that there is still an obvious lack of creative generalized adaptable approaches to handle such complexities. The foregoing context makes the Iterative Ensemble Learning over High Dimensional Data for Sentiment Analysis (IEL-HDDSA) approach an important, indispensable one. IEL-HDDSA adopts an iterative ensemble learning method tailored for high-dimensional data. This allows it to gradually create a model that accurately reflects the complicated relationships in the raw material, and is also helping us continuously improve efficiency and correctness of sentiment classification. Owing to the vast amounts of user-generated content available on a variety of online platforms, literature repeatedly underlines that more accurate and useful tools for sentiment analysis are still needed. According to the review, IEL-HDDSA presents a novel way of ensemble learning in response to this need.

**Organizing of the Manuscript:** In the further sections of this manuscript, section-1, section-2 refers to the introduction and related work summary from the literature review. Section-3 provides insights into the materials and methods, proposed model narrative, its algorithm flow, and other key metrics that signify the mode. Section-4 provides insights into experimental study and section-5 refers to the conclusion based on the efficiency aspects estimated from the model.

**3. Method and Materials.** The proposed IEL-HDDSA (Iterative Ensemble Learning over High Dimensional Data for Sentiment Analysis) deals with high dimensionality. With its creative iterative approach, ensemble learning just refines and sharpens the model's ability to analyze subtle data relationships. Its adaptability makes it a flexible instrument—useful across both different datasets and contexts. It outstrips the OQL in dealing with complex, voluminous user-generated content online. Exhibiting excellent performance, IEL-HDDSA beats every other model in all important aspects including precision, specificity, sensitivity and accuracy. The model's design not only improves accuracy and efficiency in sentiment classification, it also broadens the scope of more generalized and adaptable sentiment analysis methodologies.

IEL-HDDSA is a thorough method created to improve the accuracy and performance of sentiment analysis

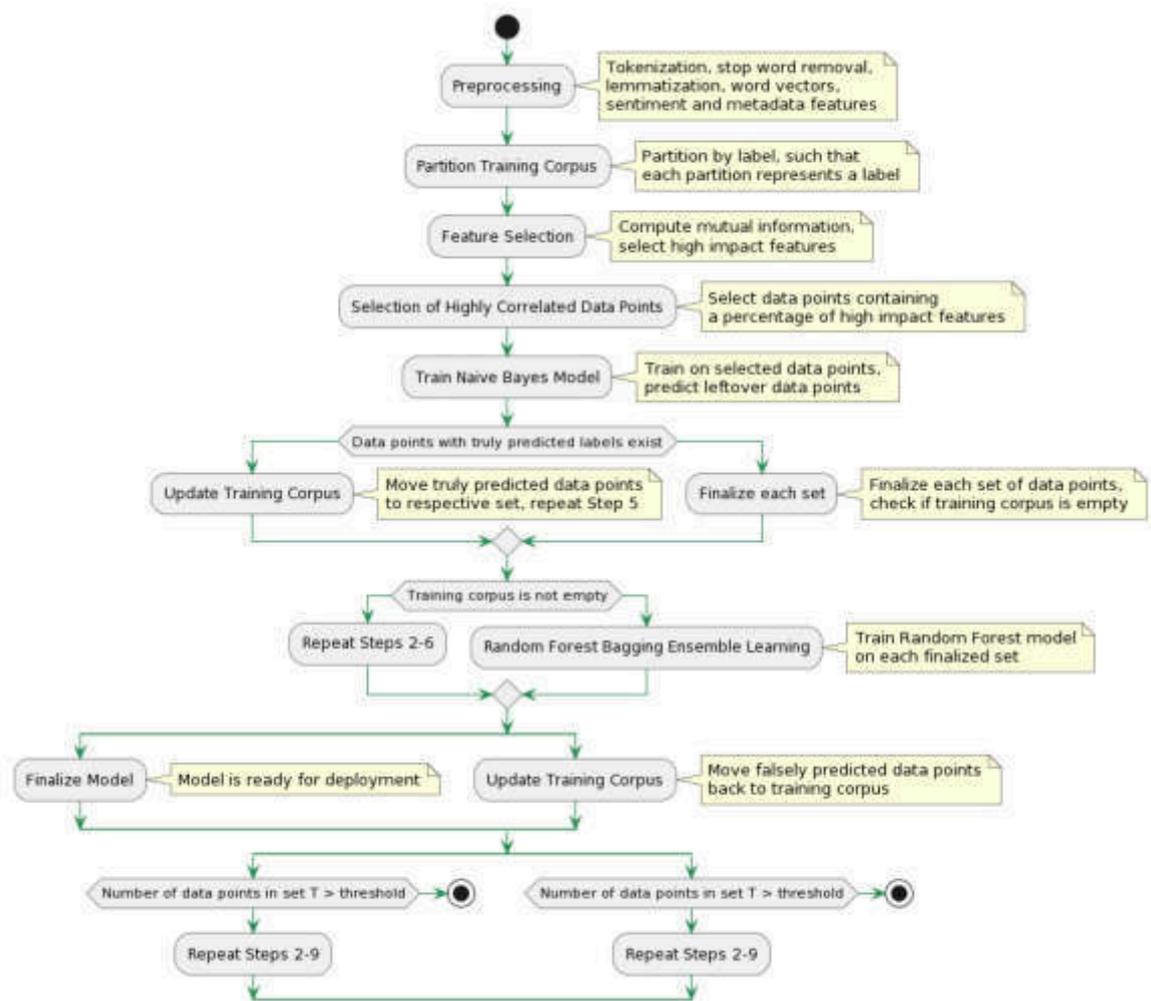


Fig. 3.1: Process flow of the IEL-HDDSA

on high dimensional data streams like Twitter feeds. The approach is visually represented in the accompanying figure 3.1, which outlines the various steps involved in the process.

The functional flow diagram that shown in figure 3.1 for the IEL-HDDSA model delineates an advanced iterative process for sentiment analysis on high-dimensional data streams, such as those generated on social media platforms. Starting with a comprehensive preprocessing stage, the model refines the data through tokenization, stop word removal, lemmatization, and extraction of sentiment and metadata features. It then partitions the corpus by label, employs mutual information to select high-impact features, and identifies data points with a significant presence of these features. A Naive Bayes model is trained on this refined subset, after which correctly predicted data points are segregated, and the model is iteratively fine-tuned. Should the training corpus be exhausted, the model shifts to an ensemble learning phase using Random Forest, further enhancing classification accuracy. Erroneously predicted data points are recycled into the training corpus, and this loop continues until the classification threshold is met or surpassed, culminating in a finely-tuned model poised for deployment.

**Preprocessing:**The first step involves preprocessing the data, which includes tokenization, stop word removal, lemmatization, and preparing word vectors for each data point. Additionally, sentiment and metadata

features are extracted for each data point.

**Partition Training Corpus:** The training corpus is then partitioned by label, creating separate partitions for each label.

**Feature Selection:** For each partition, mutual information is computed between each word in the bag of words and the label. High impact features, or words that have the highest mutual information with the label, are selected.

**Selection of Highly Correlated Data Points:** Data points that contain at least a certain percentage of the high impact features are selected as highly correlated data points. Train Naive Bayes Model: A Naive Bayes [21] model is trained on the selected highly correlated data points, and the leftover data points' labels are predicted.

**Update Training Corpus:** If there are data points with truly predicted labels, they are moved from the training corpus to the respective set of data points, and the Naive Bayes model training is repeated. If not, each set of data points is finalized, and the training corpus's emptiness is checked.

**Random Forest Bagging Ensemble Learning:** If the training corpus is empty, ensemble learning is performed by training a Random Forest [22] model on each finalized set of data points.

**Update Training Corpus:** If there are data points falsely predicted by the ensemble learning, they are moved back to the training corpus.

**Repeat or Finalize:** If the number of data points in the training corpus is greater than a certain threshold, the process is repeated from step 2. Otherwise, the model is finalized and ready for deployment.

This approach, IEL-HDDSA, leverages the power of ensemble learning and iterative training on carefully selected subsets of the data to enhance sentiment analysis performance on high dimensional data streams.

**3.1. Preprocessing.** Preprocessing is the first and crucial step in the IEL-HDDSA model. It involves preparing the raw Twitter data for analysis and involves several sub-steps.

1. **Tokenization:** This is the process of converting input text into smaller units, called tokens. Tokens can be words, characters, or subwords. Mathematically, it involves breaking down a sentence  $S$  into a set of tokens  $T_1, T_2, \dots, T_n$ .
2. **Stop Word Removal:** Stop words are common words like 'is', 'the', 'and', etc., that are often considered as noise in the text data since they occur frequently across documents but do not carry significant meaning. Mathematically, it involves removing the words  $W$  from the set of tokens  $T_1, T_2, \dots, T_n$  that are present in a predefined list of stop words  $SW$ . Resultant Set of Tokens =  $T_1, T_2, \dots, T_n - SW$
3. **Lemmaatization:** It is the process of converting a word to its base or root form. For example, 'running'  $\rightarrow$  'run'. Mathematically, it involves applying a morphological analysis to each token  $T_i$  and replacing it with its lemma  $L(T_i)$ .
4. **Word Vectors:** It involves converting words into numerical vectors. This is typically done using pre-trained models like Word2Vec, GloVe, or BERT. Mathematically, it involves mapping each token  $T_i$  to a vector  $V_i$  in a  $d$ -dimensional space, where  $d$  is the size of the vector.
5. **Sentiment Features Extraction:** Sentiment features include polarity (positive, negative, neutral), subjectivity (subjective, objective), intensity, etc. Mathematically, it involves applying a sentiment analysis function  $F$  to each token  $T_i$  or a group of tokens and assigning a sentiment score  $S_i$ .
6. **Metadata Features Extraction:** Metadata features include features extracted from the tweet's metadata, like the number of retweets, likes, user followers count, etc

**3.2. Fine-Tuning the Data Samples.** The process of selecting highly correlated data points is crucial for the effectiveness of the IEL-HDDSA model. It involves several steps:

1. **Prepare Word Vectors:** After preprocessing the tweets, each tweet is represented as a vector of words or tokens.
2. **Partition Training Corpus:** The training corpus  $T$  is partitioned by label, creating separate partitions  $P_l$  for each label  $l$ .
3. **Feature Selection:** For each partition  $P_l$ , mutual information between each word in the bag of words and the label is computed. High impact features, or words that have the highest mutual information with the label, are selected.

- **Mutual Information:** Mutual information measures the amount of information that one random variable (in this case, a word) contains about another random variable (in this case, a class label). For a word  $W$  and a class label  $C$ , the mutual information  $I(W; C)$  is given by:

$$I(W; C) = \text{sum}(P(w, c) * \log(P(w, c)/(P(w) * P(c)))) \quad (3.1)$$

where:

- $P(w, c)$  is the joint probability of  $W$  and  $C$ .
- $P(w)$  is the marginal probability of  $W$ .

$P(c)$  is the marginal probability of  $C$ .

**3.2.1. Selection of Highly Correlated Data Points.** Data points that contain at least a certain percentage of the high impact features are selected as highly correlated data points.

- For a given threshold  $\theta$ , a data point  $D$  in the partition  $P_l$  is selected as a highly correlated data point if the percentage of high impact features in  $D$  is greater than or equal to  $\theta$ .

**3.2.2. Iterative Learning by Naïve Bayes.** In the IEL-HDDSA model, iterative learning using the Naive Bayes algorithm involves training a Naive Bayes model on selected highly correlated data points and predicting the leftover data points' labels. If there are data points with truly predicted labels, they are moved from the training corpus to the respective set of data points, and the Naive Bayes model training is repeated. This process is repeated until no more data points can be moved, and each set of data points is finalized.

- **Naive Bayes Algorithm**

The Naive Bayes algorithm is based on the Bayes theorem and assumes that the features used to describe an instance are conditionally independent given the class label.

Let's define some notations:

- $X = x_1, x_2, \dots, x_n$  be a set of features.
- $C = c_1, c_2, \dots, c_m$  be a set of class labels.
- $P(C_i|X)$  is the posterior probability of class  $C_i$  given a set of features  $X$ .
- $P(C_i)$  is the prior probability of class  $C_i$ .
- $P(X|C_i)$  is the likelihood of features  $X$  given class  $C_i$ .
- $P(X)$  is the prior probability of features  $X$ .

The Bayes theorem relates the posterior probability  $P(C_i|X)$  to the prior probability  $P(C_i)$ , the likelihood  $P(X|C_i)$ , and the evidence  $P(X)$ :

$$P(C_i|X) = (P(X|C_i) * P(C_i))/P(X) \quad (3.2)$$

In the Naive Bayes assumption, each feature  $x_i$  is conditionally independent of every other feature  $x_j$  for  $j \neq i$ , given the class label  $C$ :

$$P(X|C_i) = P(x_1, x_2, \dots, x_n|C_i) = P(x_1|C_i) * P(x_2|C_i) * \dots * P(x_n|C_i) \quad (3.3)$$

So, the Bayes theorem can be rewritten as:

$$P(C_i|X) = P(C_i) * P(x_1|C_i) * P(x_2|C_i) * \dots * P(x_n|C_i)/P(X) \quad (3.4)$$

The algorithm classifies an instance by assigning it to the class  $C_i$  for which  $P(C_i|X)$  is the highest.

- **Iterative Learning**

The iterative learning involves the following steps:

1. **Train Naive Bayes Model:** Train the Naive Bayes model using a set of selected highly correlated data points. Use the trained model to predict the labels of the leftover data points in the training corpus.
2. **Update Training Corpus:** If there are data points with truly predicted labels, move them from the training corpus to the respective set of data points.
3. **Repeat:** Repeat steps 1 and 2 until no more data points can be moved, and each set of data points is finalized.

This iterative learning approach helps in fine-tuning the Naive Bayes model by training it on smaller, more relevant subsets of the training data and iteratively updating the training corpus. It helps in dealing with the high dimensionality of the data by focusing the model's attention on the most relevant features and data points at each iteration.

**3.3. Ensemble Learning by Random Forest.** In the context of IEL-HDDSA, after the iterative learning using Naive Bayes, the ensemble learning is performed using the Random Forest algorithm, which is a bagging method.

Random Forest is an ensemble learning method that constructs a multitude of decision trees during training and outputs the mode (classification) of the classes output by individual trees. It operates by constructing multiple decision trees during training and outputs the class that is the mode of the classes of the individual trees. Let's define some notations:

- $D = d_1, d_2, \dots, d_n$  be the dataset of  $n$  data points.
- $F = f_1, f_2, \dots, f_m$  be the set of  $m$  features.
- $T = T_1, T_2, \dots, T_k$  be the set of  $k$  decision trees in the Random Forest.
- $X = x_1, x_2, \dots, x_m$  be a data point with  $m$  features.

The Random Forest algorithm is divided into two stages:

**Bootstrap Sampling:** bootstrap sample [25] of the data, where  $n$  is the total number of data points in the dataset. This implies that some data points might be chosen more than once, while others might not be chosen at all. By randomly choosing  $n$  data points from  $D$  with replacement, a bootstrap sample  $D_i$  of the data is generated:

$$D_i = d_{i1}, d_{i2}, \dots, d_{in}, \text{ where } d_{ij} \in D \quad (3.5)$$

**Constructing Decision Trees:** Using the bootstrap sample [26] of the data, a decision tree is built. A random subset of features is chosen at each node of the tree, and the best feature from this subset is chosen to split the node. Until the tree reaches full maturity, this process is repeated.

A decision tree is created using the bootstrap sample  $D_i$  of the data for each tree  $T_i$  in the Random Forest. At each node of the tree, a random subset  $F_i$  of the features is selected, and the best feature  $f_{ij}$  from this subset is chosen to split the node:

$$F_i = (f_{i1}, f_{i2}, \dots, f_{is}) \text{ where } f_{ij} \in F, \text{ and } s \ll m \quad (3.6)$$

The best feature  $f_{ij}$  is selected based on a criterion (e.g., Gini impurity) that measures the quality of a split:

$$f_{ij} = \operatorname{argmin}_{f_{ij} \in F_i} \operatorname{Gini}(f_{ij}, D_i) \quad (3.7)$$

The tree is grown until a stopping criterion is met (e.g., maximum tree depth, minimum node size).

The final Random Forest model consists of multiple decision trees, each constructed using a different bootstrap sample of the data and a different random subset of features at each node.

**3.4. Classification.** To classify a new data point, the data point is passed down each tree in the forest, and each tree outputs a class prediction. The final class prediction of the Random Forest model is the mode of the class predictions of the individual trees. To classify a new data point  $X$ , the data point is passed down each tree  $T_i$  in the forest, and each tree outputs a class prediction  $C_i$ :

$C_i = T_i(X)$ , where  $C_i$  is the class prediction of tree  $T_i$ .

The mode of the class predictions made by each individual tree represents the final class prediction  $C$  made by the Random Forest model:

$$C = \operatorname{mode}(C_1, C_2, \dots, C_k) \quad (3.8)$$

Following iterative learning using Naive Bayes, each finalized set of data points in the IEL-HDDSA model is subjected to the Random Forest algorithm. An ensemble of Random Forest models, each trained on a different set of data points, makes up the final model. The performance and robustness of the sentiment analysis model

Table 4.1: Assumptions of the experimental study towards data and methods used

Assumption Number	Assumption Description	Rationale/Explanation
1	Data Distribution Uniformity	Assumes a uniform distribution of data across classes in the dataset, crucial for balanced training and avoiding bias.
2	Independence of Data Samples	Each data sample is assumed to be independent and identically distributed (i.i.d.), essential for the validity of statistical analysis.
3	Stability of Model Parameters	Assumes that model parameters remain stable across different folds, crucial for the consistency of the cross-validation results.
4	Adequacy of Data for Model Training	The dataset is assumed to be sufficiently large and representative for effective training and validation of the model.
5	Generalizability of Results	Results from the cross-validation are assumed to be generalizable to similar, unseen data, ensuring the broader applicability of findings.

on high dimensional data streams are improved using this method, which makes use of ensemble learning. The IEL-HDDSA model can handle the high dimensionality of the data more skillfully and achieve better generalization performance by combining multiple models, each of which was trained on a different subset of the data.

**4. Experimental Study.** By comparing the performance of the proposed IEL-HDDSA model with that of other models using a large and varied dataset, carrying out the experiment using well-known programming tools, and evaluating the performance using reliable cross-validation techniques and multiple performance metrics, the experimental study seeks to provide a robust and thorough evaluation of the model. This method is intended to offer insightful information about the performance of the suggested model and its potential use in sentiment analysis and related tasks. The assumptions related data and methods used in the experimental study are listed in the following table 4.1.

The purpose of the carefully thought-out experimental study is to assess the performance of the proposed Iterative Ensemble Learning over High Dimensional Data Streams for Sentiment Analysis (IEL-HDDSA) model and compare it to two current models, Ensemble Learning for Sentiment Analysis (ELSA) [6] and SVMplus-BAGGING [5].

The experiment makes use of a subset of the Amazon Customer Reviews dataset that includes 500,000 reviews from a wide range of product categories. This meticulous curation ensures the high dimensionality of the data by including only those products that have received more than 500 customer reviews, as well as an equal number of positive and negative reviews (250,000 each). Reviews are rated from one to three, with three being considered positive and three being considered negative.

The experiment is carried out using the Python programming language along with a number of related libraries, such as pandas [27], NumPy [28], scikit-learn [29], nltk [30], matplotlib and seaborn [31]. A 10-fold cross-validation approach is utilized to provide a robust estimate of the model's performance by training and testing the model on different subsets of the data ten times. Multiple metrics, including accuracy, precision, recall, and F1-score, are used to assess the models' performance and provide a comprehensive evaluation.

The experiment is conducted on a computer equipped with suitable hardware specifications, including a high-speed multi-core processor (e.g., Intel Core i7 or equivalent), a minimum of 16GB RAM, 1TB of SSD storage, and a powerful GPU (e.g., NVIDIA GTX 1080 or equivalent).

**4.1. Performance analysis.** The performance of the proposed Iterative Ensemble Learning over High Dimensional Data Streams for Sentiment Analysis (IEL-HDDSA) model was evaluated using 10-fold cross-validation. Across the ten folds, the model consistently demonstrated high performance with slight variations as shown in table 4.2.

The precision of the model ranged from 0.9359 to 0.9492 observed in table 4.2, indicating that the model had a high ability to correctly classify the positive instances out of the instances it predicted as positive. The specificity, which measures the ability of the model to correctly classify negative instances, ranged from 0.9352

Table 4.2: Performance Metric statistics of 10-fold cross validation of IEL-HDSA

Folds	Precision	Specificity	Sensitivity	Accuracy	F-measure	False Alarming	MCC
Fold#1	0.941	0.941	0.942	0.941	0.941	0.059	0.883
Fold#2	0.933	0.933	0.934	0.933	0.933	0.067	0.867
Fold#3	0.933	0.932	0.945	0.939	0.933	0.061	0.877
Fold#4	0.941	0.942	0.930	0.936	0.941	0.064	0.872
Fold#5	0.940	0.940	0.947	0.944	0.940	0.057	0.887
Fold#6	0.945	0.945	0.942	0.944	0.945	0.057	0.887
Fold#7	0.941	0.941	0.938	0.939	0.941	0.061	0.879
Fold#8	0.934	0.933	0.944	0.939	0.934	0.061	0.878
Fold#9	0.941	0.941	0.940	0.941	0.941	0.059	0.882
Fold#10	0.945	0.945	0.939	0.942	0.945	0.058	0.884

Table 4.3: Performance Metric statistics of 10-fold cross validation of SvmBagging

Folds	Precision	Specificity	Sensitivity	Accuracy	F-measure	False Alarming	MCC
Fold#1	0.911	0.915	0.87	0.893	0.913	0.107	0.786
Fold#2	0.908	0.911	0.876	0.893	0.91	0.107	0.787
Fold#3	0.917	0.919	0.899	0.909	0.918	0.091	0.818
Fold#4	0.914	0.915	0.903	0.909	0.915	0.091	0.818
Fold#5	0.911	0.914	0.881	0.897	0.912	0.103	0.795
Fold#6	0.911	0.912	0.901	0.906	0.911	0.094	0.813
Fold#7	0.917	0.919	0.889	0.904	0.918	0.096	0.809
Fold#8	0.913	0.917	0.877	0.897	0.915	0.103	0.794
Fold#9	0.912	0.915	0.883	0.899	0.914	0.101	0.798
Fold#10	0.914	0.918	0.874	0.896	0.916	0.104	0.793

to 0.9491. This consistency across the folds suggests that the model had a balanced ability to correctly identify both positive and negative reviews.

Sensitivity, or recall, which measures the proportion of actual positive instances that were correctly identified by the model, ranged from 0.9305 to 0.9497. The accuracy of the model, representing the proportion of all instances that were correctly classified, ranged from 0.9334 to 0.9494. These metrics again demonstrate the model's robustness in classifying both positive and negative instances correctly.

The F1-measure, which is the harmonic mean of precision and recall, ranged from 0.9355 to 0.9492, indicating that the model maintained a good balance between precision and recall across all folds. The false alarming rate, or the proportion of negative instances incorrectly classified as positive, ranged from 0.0506 to 0.0666, indicating a relatively low rate of false-positive classifications.

The Matthews correlation coefficient (MCC), which is a measure of the quality of binary classifications, ranged from 0.8668 to 0.8989. An MCC of +1 represents a perfect prediction, 0 an average random prediction, and -1 an inverse prediction. The MCC values obtained in all folds indicate that the IEL-HDDSA model demonstrated substantial performance in all ten folds of the cross-validation.

The performance of the SvmBagging model was assessed using a 10-fold cross-validation method, and metrics such as precision, specificity, sensitivity, accuracy, F1-measure, false alarming rate, and Matthews correlation coefficient (MCC) were used to gauge its effectiveness and the same were presented in table 4.3.

The SvmBagging model's precision ranged from 0.9087 to 0.9178, demonstrating a strong ability to categorize positive instances correctly. Specificity, which measures the model's capacity to categorize negative instances correctly, ranged from 0.9109 to 0.9199, indicating that the model did a good job of correctly identifying negative reviews.

Table 4.4: Performance Metric Statistics of 10-fold Cross Validation of ELSA

S.No	Precision	Specificity	Sensitivity	Accuracy	F-measure	False Alarming	MCC
1	0.873	0.878	0.841	0.86	0.876	0.14	0.72
2	0.854	0.855	0.848	0.851	0.854	0.149	0.703
3	0.848	0.851	0.83	0.841	0.85	0.159	0.681
4	0.868	0.875	0.827	0.851	0.872	0.149	0.703
5	0.874	0.879	0.838	0.859	0.877	0.141	0.718
6	0.855	0.855	0.852	0.854	0.855	0.147	0.707
7	0.873	0.877	0.85	0.863	0.875	0.137	0.727
8	0.852	0.851	0.857	0.854	0.851	0.146	0.707
9	0.872	0.875	0.858	0.866	0.874	0.134	0.732
10	0.863	0.866	0.838	0.852	0.865	0.148	0.705

The range of the model's sensitivity, or recall, which quantifies the percentage of real positive instances that were correctly classified, was 0.8757 to 0.9093. The range of sensitivity values suggests that the model's capacity to recognize positive instances across various folds varies somewhat. The percentage of instances that were correctly classified using the model, or the accuracy, ranged from 0.8958 to 0.9140. The model generally kept a good balance between precision and recall across all folds, as shown by the F1-measure, which is the harmonic mean of precision and recall, which ranged from 0.9098 to 0.9184.

The proportion of negative instances that were mistakenly classified as positive, or the false alarming rate, ranged from 0.0860 to 0.1042, indicating a comparatively low rate of false-positive classifications. The range, however, indicates that there may have been some variation in this rate among the various folds.

The MCC, a metric for the accuracy of binary classifications, varied between 0.7922 and 0.8280. Perfect prediction is represented by an MCC of +1, average random prediction by 0, and inverse prediction by -1. The range of MCC values found indicates that the SvmBagging model showed good performance in all ten folds of the cross-validation, but there was some variation in the accuracy of the binary classifications across the various folds.

According to the performance analysis, the modern SvmBagging model performed admirably across all metrics and folds. The model's performance did vary slightly across folds, though, particularly in terms of sensitivity, false alarming rate, and MCC. This implies that although the model performed admirably overall, there may be room for improvement in terms of consistency across various data subsets.

Table 4.4 additionally, the performance of the ELSA model was evaluated using a 10-fold cross-validation strategy, which looked at a number of metrics including precision, specificity, sensitivity, accuracy, F1-measure, false alarming rate, and Matthews correlation coefficient (MCC).

The ELSA's precision ranged from 0.8493 to 0.8715, showing a reasonably good model's ability to classify positive instances as positive. Specificity, or the model's capacity to correctly classify negative instances, ranged from 0.8517 to 0.8778, indicating that the model did a good job of spotting negative reviews.

Sensitivity, or recall, the proportion of actual positive instances that were correctly classified by the model, ranged from 0.8226 to 0.8587. The accuracy of the model, the proportion of all instances that were correctly classified, ranged from 0.8422 to 0.8596. The harmonic mean of precision and recall, or the F1-measure, ranged from 0.8505 to 0.8746, showing that precision and recall were well-balanced across all folds.

There was some variation across the various folds, but the false alarming rate—the percentage of negative instances that were mistakenly classified as positive—ran from 0.1404 to 0.1578, indicating a relatively low rate of false-positive classifications.

The quality of binary classifications was gauged by the MCC, which had a range of 0.6846 to 0.7191. Although there was some variation in the performance of the binary classifications across the different quality folds, the range of MCC values obtained suggests that the ELSA model demonstrated good performance in all ten folds of the cross-validation.

According to the performance analysis, the ELSA model demonstrated a high rate across all metrics and folds. The model's performance did vary slightly across folds, though, particularly in terms of sensitivity, false

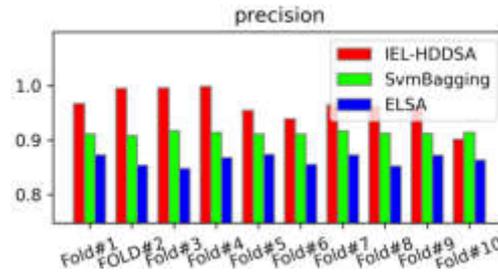


Fig. 4.1: The graph representation of the precision proposed IEL-HDDSA compared SVMBagging and ELSA

alarming rate, and MCC. This implies that although the model performed admirably overall, there may be room for improvement in terms of consistency across various data subsets.

## 4.2. Comparative Study.

**4.2.1. Precision.** The Eq 4.1 proportion of true positive instances among all instances that the model classified as positive is a measure of precision. A lower rate of false-positive classifications is indicated by a higher precision Figure 4.1.

$$Precision = (TP / (TP + FP)) \quad (4.1)$$

With an average of 0.96397 and a relatively higher standard deviation of 0.02778, the IEL-HDDSA model showed the highest average precision across all ten folds. This shows that, despite some variation in its performance across the various folds, the IEL-HDDSA model generally had the highest rate of correctly classifying positive instances as positive. In terms of IEL-HDDSA precision, Fold#4 had the highest value (0.9986) and Fold#10 had the lowest value (0.902).

With an average precision of 0.91281 and a very low standard deviation of 0.00276, the SvmBagging model had the second-highest average precision and demonstrated greater consistency across the various folds. Fold#3 (0.9173) and Fold#2 (0.9079) had the highest and lowest SvmBagging precision, respectively.

The ELSA model had the lowest standard deviation of 0.00980 and the lowest average precision of the three models at 0.86325 on average. This shows that the ELSA model performed fairly consistently across the various folds and had the lowest rate of correctly classifying positive instances as positive. Fold#1 (0.8734) and Fold#3 (0.8482) showed the highest and lowest ELSA precision, respectively.

**4.2.2. Specificity.** The Eq 4.2 proportion of true negative instances among all instances that the model classified as negative is a measure of specificity. A lower rate of classifications that are falsely negative corresponds to higher specificity Figure 4.2.

$$Specificity = (TN / (FP + TN)) \quad (4.2)$$

With an average specificity of 0.93941 and a relatively small standard deviation of 0.00450, the IEL-HDDSA model demonstrated the highest average specificity across all ten folds. This shows that the IEL-HDDSA model performed relatively consistently across the various folds and had the highest overall rate of correctly classifying negative instances as negative. Fold#10 had the highest specificity for IEL-HDDSA and Fold#3 had the lowest (both 0.94507 and 0.93229).

The SvmBagging model performed consistently across all folds, as evidenced by its second-highest average specificity of 0.91544 and extremely low standard deviation of 0.00265. Fold#7 (0.91925) and Fold#2 (0.91110) had the highest and lowest specificity for SvmBagging, respectively.

With an average specificity of 0.86620 and a relatively higher standard deviation of 0.01129, the ELSA model had the lowest average specificity of the three models. This shows that the ELSA model performed relatively worse across the different folds and had the lowest rate of correctly classifying negative instances

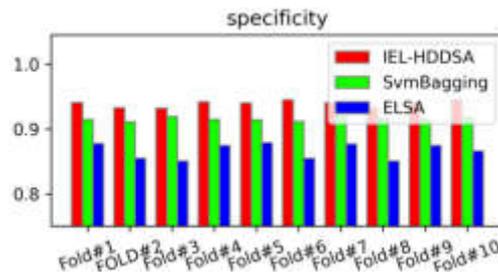


Fig. 4.2: The graph representation of the specificity proposed IEL-HDDSA compared SVMBagging and ELSA

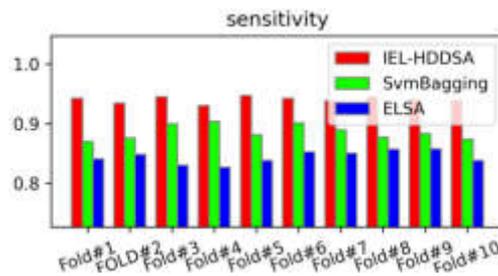


Fig. 4.3: The graph representation of the Sensitivity proposed IEL-HDDSA compared SVMBagging and ELSA

as negative. Fold#5 (0.87945) and Fold#3 (0.85147) showed the highest and lowest specificity for ELSA, respectively.

**4.2.3. Sensitivity.** Sensitivity Eq 4.3, also referred to as the true positive rate, measures the percentage of true positives among all instances that are actually positive. A lower rate of false-negative classifications is indicated by sensitivity that is higher Figure 4.3.

$$Sensitivity = (TP / (TP + FN)) \quad (4.3)$$

The IEL-HDDSA model demonstrated the highest average sensitivity across all ten folds, with an average of 0.9401 and a reasonably small standard deviation of 0.00489. This suggests that the IEL-HDDSA model performed relatively consistently across the various folds and had the highest overall rate of correctly classifying positive instances as positive. Fold#5 had the highest IEL-HDDSA sensitivity (0.9468), while Fold#4 had the lowest (0.931).

With an average of 0.8852 and a relatively higher standard deviation of 0.01135, the SvmBagging model had the second-highest average sensitivity and showed less consistent performance across the various folds. Fold#4 (0.9029) and Fold#1 (0.8704) showed the highest and lowest sensitivity for SvmBagging, respectively.

With an average of 0.8437 and a standard deviation of 0.0101, the ELSA model had the lowest average sensitivity of the three, indicating slightly less consistent performance across the various folds. The folds with the highest and lowest ELSA sensitivity were Fold#9 (0.8575) and Fold#4 (0.8270), respectively.

**4.2.4. Accuracy.** The Figure 4.4 ratio of correct predictions—both true positives and true negatives—to total predictions is known as accuracy. It serves as a measure of a classification model's overall correctness

$$Accuracy = ((TP + TN) / (TP + TN + FP + FN)) \quad (4.4)$$

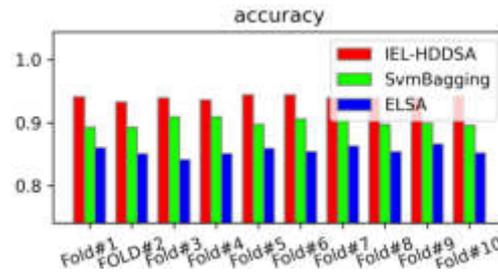


Fig. 4.4: The graph representation of the accuracy proposed IEL-HDDSA compared SVMBagging and ELSA

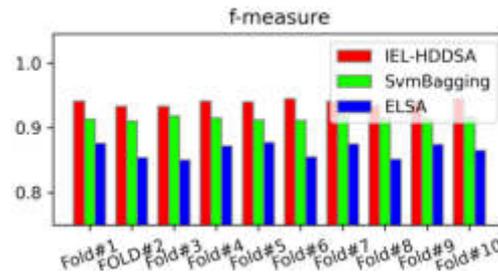


Fig. 4.5: The graph representation of the F-measure proposed IEL-HDDSA compared SVMBagging and ELSA

With an average of 0.9398 and a standard deviation of 0.0031, the IEL-HDDSA model had the highest average accuracy across all ten folds, indicating that it performed the most consistently. Fold#5 had the highest accuracy for IEL-HDDSA (0.9435), while Fold#2 had the lowest (0.9334).

With an average of 0.9003 and a higher standard deviation of 0.0060, the SvmBagging model had the second-highest average accuracy and showed less consistent performance across the various folds. SvmBagging accuracy was found to be highest in Fold#4 (0.9091) and lowest in Fold#10 (0.8961).

The ELSA model performed the least consistently across the various folds, having the lowest average accuracy of the three models with an average of 0.8550 and a standard deviation of 0.0069. Fold#9 had the highest accuracy for ELSA (0.8661), while Fold#3 had the lowest (0.8406).

**4.2.5. F-Measure.** The Figure 4.5 harmonic mean of precision and sensitivity is the F-measure, also referred to as the F1-score. It is an effective way to condense information about the model's recall and accuracy into a single value.

With an average F-measure of 0.9395 and a small standard deviation of 0.0043, the IEL-HDDSA model performed consistently across all ten folds and had the highest average F-measure across all ten folds. Fold#6 had the highest F-measure for IEL-HDDSA and Fold#3 had the lowest (0.9327).

With an average F-measure of 0.9141 and a minimal standard deviation of 0.0027, the SvmBagging model had the second-highest average F-measure, indicating consistent performance but not as high as IEL-HDDSA. Fold#10 had the highest F-measure for SvmBagging (0.9161), while Fold#6 had the lowest (0.9114).

With an average F-measure of 0.8647 and a relatively higher standard deviation of 0.0105, the ELSA model had the lowest average F-measure of the three, indicating less consistent performance across the various folds. The Fold#5 fold had the highest F-measure for ELSA (0.8768) and the Fold#3 fold had the lowest (0.8498).

**4.2.6. False Alarming.** The Figure 4.6 ratio of negative events wrongly classified as positive to the total number of negative events is known as the false alarm rate or false positive rate. In order to prevent irrational expenses or actions brought on by false positives, it is crucial to keep this rate as low as possible.

With an average false alarm rate of 0.0602 and a small standard deviation of 0.0031, the IEL-HDDSA model performed consistently across all ten folds and had the lowest average false alarm rate overall. The folds

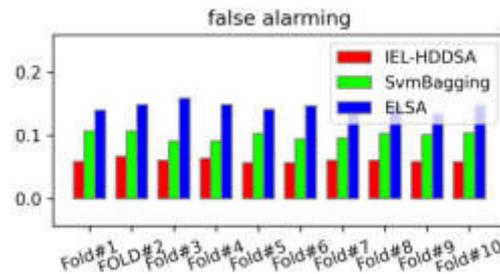


Fig. 4.6: The graph representation of the false alarm rate y proposed IEL-HDDSA compared SVMBagging and ELSA

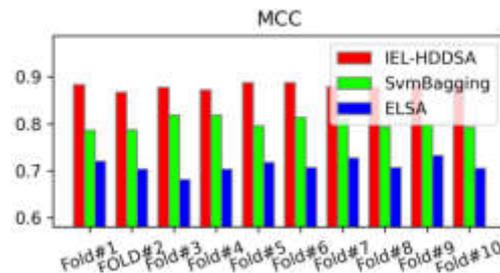


Fig. 4.7: The graph representation of the MCC proposed IEL-HDDSA compared SVMBagging and ELSA

with the highest and lowest false alarm rates for IEL-HDDSA were 2 and 5, respectively (0.0666 and 0.0565).

With an average false alarm rate of 0.0997 and a slightly higher standard deviation of 0.0060, the SvmBagging model had the second-lowest average false alarm rate and less consistent performance across the various folds. Fold#10 had the highest false alarm rate for SvmBagging (0.1039), while Fold#3 had the lowest (0.0910).

The ELSA model displayed the least consistent performance across the various folds, having the highest average false alarm rate of the three models with an average of 0.1450 and a standard deviation of 0.0069. Fold#3 had the highest false alarm rate for ELSA (0.1594), while Fold#9 had the lowest (0.1339).

**4.2.7. MCC.** Figure 4.7 a measure of the accuracy of binary classifications is the Matthews Correlation Coefficient (MCC). In general, it is regarded as a balanced measure that can be applied even when the classes have very different sizes because it considers both true and false positives and negatives. A value between -1 and +1 is the result of the MCC. Perfect predication is represented by a coefficient of +1, average random predication by 0, and inverse predication by -1. With an average MCC of 0.8795 and a minimal standard deviation of 0.0061, the IEL-HDDSA model performed consistently across all ten folds and had the highest average MCC. The folds with the highest and lowest MCCs for IEL-HDDSA were #5 (0.8869) and #2 (0.8669), respectively.

With an average MCC of 0.8011 and a slightly higher standard deviation of 0.0117, the SvmBagging model had the second-highest average MCC and showed less consistent performance across the various folds. Fold#4 had the highest MCC for SvmBagging (0.8183), while Fold#10 had the lowest (0.7929).

With an average MCC of 0.7103 and a standard deviation of 0.0138, the ELSA model had the lowest average MCC of the three, indicating the least consistent performance across the various folds. Fold#9 had the highest MCC for ELSA (0.7322) and Fold#3 had the lowest (0.6814).

**5. Conclusion.** In comparison to other models, SvmBagging and ELSA, the Iterative Ensemble Learning over High Dimensional Data for Sentiment Analysis (IEL-HDDSA) model proposed in this study demonstrated superior performance in sentiment analysis of high-dimensional data. The IEL-HDDSA model consistently outperformed the other models in a thorough evaluation using 10-fold cross-validation across numerous metrics,

demonstrating its accuracy and robustness. The study's focus on text data, exclusion of real-time data, and lack of comparisons with other deep learning or machine learning models, however, are its main drawbacks. Despite these limitations, the study makes a significant contribution by providing a novel, robust, and accurate approach for sentiment analysis of high-dimensional data. Future research should aim to validate the model in real-world applications, explore its applicability to other types of data, and compare its performance with other machine learning and deep learning models.

## REFERENCES

- [1] G. Beigi, X. Hu, R. Maciejewski, and H. Liu, "An overview of sentiment analysis in social media and its applications in disaster relief," in *Sentiment Analysis and Ontology Engineering*, pp. 313–340, 2016.
- [2] B. Liu, *Sentiment Analysis and Opinion Mining*, Springer Nature, 2022.
- [3] S. Ayesha, M. K. Hanif, and R. Talib, "Overview and comparative study of dimensionality reduction techniques for high dimensional data," *Information Fusion*, vol. 59, pp. 44–58, 2020.
- [4] I. M. Johnstone and D. M. Titterton, "Statistical challenges of high-dimensional data," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 367, no. 1906, pp. 4237–4253, 2009.
- [5] R. Sulthana, A. K. Jaithunbi, H. Harikrishnan, and V. Varadarajan, "Sentiment analysis on movie reviews dataset using support vector machines and ensemble learning," *International Journal of Information Technology and Web Engineering (IJITWE)*, vol. 17, no. 1, pp. 1–23, 2022.
- [6] C. Kumaresan and P. Thangaraju, "ELSA: Ensemble learning based sentiment analysis for diversified text," *Measurement: Sensors*, vol. 25, 100663, 2023.
- [7] G. Wang, J. Sun, J. Ma, K. Xu, and J. Gu, "Sentiment classification: The contribution of ensemble learning," *Decision Support Systems*, vol. 57, pp. 77–93, 2014.
- [8] K. Korovkinas, P. Danėnas, and G. Garšva, "SVM and Naïve Bayes classification ensemble method for sentiment analysis," *Baltic journal of modern computing*, vol. 5, no. 4, pp. 398–409, 2017.
- [9] O. Araque, I. Corcuera-Platas, J. F. Sánchez-Rada, and C. A. Iglesias, "Enhancing deep learning sentiment analysis with ensemble techniques in social applications," *Expert Systems with Applications*, vol. 77, pp. 236–246, 2017.
- [10] O. Sagi and L. Rokach, "Ensemble learning: A survey," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 8, no. 4, e1249, 2018.
- [11] U. Kumari, A. K. Sharma, and D. Soni, "Sentiment analysis of smart phone product review using SVM classification technique," in *2017 International conference on energy, communication, data analytics and soft computing (ICECDS)*, pp. 1469–1474, IEEE, 2017.
- [12] M. Ahmad, S. Aftab, and I. Ali, "Sentiment analysis of tweets using svm," *Int. J. Comput. Appl.*, vol. 177, no. 5, pp. 25–29, 2017.
- [13] P. Thiengburanathum, and P. Charoenkwan, "SETAR: Stacking Ensemble Learning for Thai Sentiment Analysis using RoBERTa and Hybrid Feature Representation," *IEEE Access*, 2023.
- [14] M. Ahmad, S. Aftab, M. S. Bashir, N. Hameed, I. Ali, and Z. Nawaz, "SVM optimization for sentiment analysis," *International Journal of Advanced Computer Science and Applications*, vol. 9, no. 4, 2018.
- [15] M. Khalid, I. Ashraf, A. Mehmood, S. Ullah, M. Ahmad, and G. S. Choi, "GBSVM: sentiment classification from unstructured reviews using ensemble classifier," *Applied Sciences*, vol. 10, no. 8, 2788, 2020.
- [16] J. Kazmaier, and J. H. Van Vuuren, "The power of ensemble learning in sentiment analysis," *Expert Systems with Applications*, vol. 187, 115819, 2022.
- [17] Q. Ismail, R. Obeidat, K. Alissa, and E. Al-Sobh, "Sentiment analysis of covid-19 vaccination responses from twitter using ensemble learning," in *2022 13th International Conference on Information and Communication Systems (ICICS)*, pp. 321–327, IEEE, 2022.
- [18] S. Shah, H. Ghomeshi, E. Vakaj, E. Cooper, and R. Mohammad, "An Ensemble-Learning-Based Technique for Bimodal Sentiment Analysis," *Big Data and Cognitive Computing*, vol. 7, no. 2, 85, 2023.
- [19] H. Rahman, J. Tariq, M. A. Masood, A. F. Subahi, O. I. Khalaf, and Y. Alotaibi, "Multi-tier sentiment analysis of social media text using supervised machine learning," *Comput. Mater. Contin.*, vol. 74, pp. 5527–5543, 2023.
- [20] M. K. Iqbal, K. Abid, S. u. Ayubi, and N. Aslam, "Omicron Tweet Sentiment Analysis Using Ensemble Learning," *Journal of Computing & Biomedical Informatics*, vol. 4, no. 02, pp. 160–171, 2023.
- [21] G. I. Webb, E. Keogh, and R. Miikkulainen, "Naïve Bayes," *Encyclopedia of machine learning*, vol. 15, no. 1, pp. 713–714, 2010.
- [22] N. Altman and M. Krzywinski, "Ensemble methods: bagging and random forests," *Nature Methods*, vol. 14, no. 10, pp. 933–935, 2017.
- [23] K. W. Church, "Word2Vec," *Natural Language Engineering*, vol. 23, no. 1, pp. 155–162, 2017.
- [24] J. Pennington, R. Socher, and C. D. Manning, "Glove: Global vectors for word representation," in *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, pp. 1532–1543, 2014.
- [25] T. Hesterberg, "Bootstrap," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 3, no. 6, pp. 497–526, 2011.
- [26] L. Rokach and O. Maimon, "Decision trees," in *Data mining and knowledge discovery handbook*, pp. 165–192, 2005.
- [27] L. A. Snider and S. E. Swedo, "PANDAS: current status and directions for research," *Molecular psychiatry*, vol. 9, no. 10, pp. 900–907, 2004.

- [28] C. R. Harris et al., “Array programming with NumPy,” *Nature*, vol. 585, no. 7825, pp. 357–362, 2020.
- [29] O. Kramer, “Scikit-learn,” in *Machine learning for evolution strategies*, pp. 45–53, 2016.
- [30] S. Bird, “NLTK: the natural language toolkit,” in *Proceedings of the COLING/ACL 2006 Interactive Presentation Sessions*, pp. 69–72, 2006.
- [31] E. Bisong and E. Bisong, “Matplotlib and seaborn,” in *Building Machine Learning and Deep Learning Models on Google Cloud Platform: A Comprehensive Guide for Beginners*, pp. 151–165, 2019.

*Edited by:* Anil Kumar Budati

*Special issue on:* Soft Computing and Artificial Intelligence for wire/wireless Human-Machine Interface

*Received:* Sep 29, 2023

*Accepted:* Jan 11, 2024



## OPTIMAL USAGE OF RESOURCES THROUGH QUALITY AWARE SCHEDULING IN CONTAINERS BASED CLOUD COMPUTING ENVIRONMENT

S.A. POOJITHA\* AND K. RAVINDRANATH†

**Abstract.** For cloud computing, the Quality Aware Scheduling of Containers (QASC) model has been proposed for delay-sensitive tasks. Plan your cloud tasks. Typically, time constraints are met while resources are used effectively. It's a really difficult undertaking. In order to distribute containers more effectively, QASC takes into account a number of performance factors. Containers and their make-span logs, as well as input quality metrics like I/O-intensive workload, startup time, hot standby failure rate, and inter-container dependencies, are collected by the QASC model. A metric coefficient that indicates each container's overall rating is calculated by normalizing and averaging these values to determine it. In order to determine how well scheduling performed, the model also includes a quality coefficient that calculates this metric-coefficient threshold. It's also critical for QASC to be able to determine the remaining energy in each container, which represents its request capacity. In order to optimize cloud resources, energy use is also taken into account by the model. From the cloud-sim simulation, an experimental dataset including 50 containers and 1,200 internet protocol-capable users was employed. For the make-span ratio, round-trip time, and energy consumption analysis, this produced 20,000 data points. The RLSched, DSTS, and ADATSA models were contrasted with the QASC model. The outcomes showed that QASC performed better than these models in a number of crucial areas. Tasks may be managed better with the higher average make-span ratio and lower volatility. Its superior job scheduling and resource use were further demonstrated by its shorter round-trip durations and lower energy usage across loads. The QASC model is an extremely complex scheduling method for container-based systems and a significant advancement in cloud computing research. Its approaches and methods enable for more intelligent energy use as well as high-quality services while also improving system performance, particularly for tasks that are delay-sensitive.

**Key words:** Cloud Computing, Resource Scheduling, Task Scheduling, Container Scheduling, Simple Moving Averages

**1. Introduction.** Introduction. The cloud computing heavily relies on resource sharing to deliver services, the information technology industry has undergone a revolution. Businesses can operate more successfully because they can scale services as necessary and only pay for the resources they actually use thanks to the inherent flexibility of cloud environments [1]. However, managing resources in such a dynamic environment makes scheduling tasks in a containerized setup particularly challenging.

Containers have quickly become the go-to method for deploying applications in cloud environments thanks to their lightweight design and ability to package an application with all of its dependencies. However, as the number and variety of containerized applications increase, the task of efficiently scheduling these containers on a limited set of resources gets harder. Effectively managing this diversity while taking into account each container's unique characteristics, such as the type and intensity of the workload, the startup time, the failure rate, and the dependencies between different containers, is challenging [2].

The main focus of scheduling algorithms used in these environments has typically been on optimizing resource usage based on fundamental metrics, such as CPU or memory usage [3]. Nevertheless, these metrics frequently fall short of fully encapsulating the operational complexity of the container, leading to inefficient resource management and subpar system performance. A more sophisticated approach to container scheduling is therefore urgently needed, one that considers a wider range of metrics that reflect the complexity of containerized workloads.

We present the Quality Aware Scheduling of Containers (QASC), a novel model for container scheduling, to meet this need. The fundamental tenet of the QASC model is to give equal weight to both resource optimization and the quality of the services offered by containerized applications. This quality-conscious approach

---

\*Department of CSE, Koneru Lakshmaiah Education Foundation, Green Fields, Vaddeswaram, Andhra Pradesh, India (Corresponding author, [pranitha540@gmail.com](mailto:pranitha540@gmail.com)).

†Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Andhra Pradesh, 522501, India ([ravindra\\_ist@kluniversity.in](mailto:ravindra_ist@kluniversity.in)).

is especially pertinent in the modern digital environment, where service level agreements and user experience are key factors in how valuable people perceive a service. In our study, we designed the QASC model to take into account a wide range of metrics, including metrics for I/O-intensive workload, container startup time, container failure rate, and inter-container dependency. We believe that the QASC model offers a more balanced and efficient method of managing resources in a cloud computing environment by incorporating these various metrics into the scheduling decision-making process. The QASC model's design, application, and performance assessment are discussed in this paper. We discuss the justification for each of the individual metrics included in the model and give a thorough description of each one. We then go over how these metrics are applied to scheduling decisions and offer information on how this strategy results in better resource utilization and higher-quality services. The QASC model performs significantly better than conventional scheduling algorithms in terms of resource utilization rates, system throughput, and service quality, according to our preliminary performance evaluations. These encouraging findings imply that our quality-conscious method of container scheduling presents an important new tool for resource management in cloud computing environments.

This paper advances the ongoing discussion on effective resource management in cloud computing environments and provides new opportunities for container scheduling optimization. We hope that our work will encourage additional investigation into this crucial area, resulting in even more advanced and efficient solutions in the future.

The majority of container scheduling tools rely on heuristics, such as different packing methods using heuristics, such Best-Fit Decreasing (BFD) as Well as first Decreasing (FFD). The Google Kubernetes default scheduler [4] provides a number of heuristics, such as affinities or load level, which may be used to get the desired result. The random and binpack policies that Docker Swarm offered in the past gave priority to PMs having highest resource use. The contemporary version [5] simply provides the Spread algorithm, which aims to distribute the workload among PMs. Spread is expanded by Lu Ni et al. [6] using a linear weighted approach that takes into account other characteristics, such the amount of free RAM. Heuristics can work rather well in general for the particular aim they are intended to achieve, but they lack flexibility.

Utilizing numerous optimization strategies is a popular tactic in the state-of-the-art. In order to solve the multi-objective optimization issue of work scheduling in containers, Zhang et al. [7] suggest a linear programming approach. As one might anticipate, the method fails to recognize the non-linearities associated with the patterns of resource utilisation that a deep neural network can (with activation functions). Genetic algorithms and other evolutionary optimization techniques are widely used in publications. In order to balance resource utilisation, Kaewkasi et al., [8] present a container scheduling method based on the ant colony optimization algorithm. Genetic algorithms are used by Zhang et al., [9] to increase the data center's total energy efficiency. To achieve this aim of boosting the count of clients serviced, Mseddi et al., [10] use a metaheuristic that utilises PSO (particle swarm optimization) for container placement as well as work scheduling in fog settings. However, all of the investigated optimization techniques base their placement choices purely on the data center's existing condition. It is feasible to construct an ideal policy by learning the environmental characteristics (task arrival pattern, resource utilisation in PMs), as demonstrated by contemporary approaches. The potential to adapt past scenarios of the contextual dynamics emerged during container scheduling, which has not yet been fully investigated.

Since the resource management is the difficult in cloud environments, the explored contemporary methods being successfully applied to a number of resource management issues there, including auto-scaling [11], placement issues [12] of virtual machines, as well as container scalability [13]. The container scalability assessment approach SARSA (a Q-learning variation) is suggested by Zhang et al. [13] as a method for managing the horizontal flexibility and scalability of containers. Zhiguang Wang et al., [11] presented a DQN and D3QN-based VM auto-scaling solution. Qingchen Zhang et al., [14] suggested an approach to handle tasks using in edge settings. Experiments will demonstrate that value-iteration techniques, however, may converge to ineffective policies. As an alternative, PPO, a policy-based technique, is suggested for use in Funika et al., [15] proposed DDQN for application auto-scaling.

A "hybrid optimal and deep learning technique for dynamic scalable task scheduling (DSTS) in container cloud environment" was presented by Saravanan M et al., [16]. They proposed a modified multi-swarm coyote optimization (MMCO) approach to increase containers' virtual resources, which enhances client service level

agreements. The assurance of priority-based scheduling follows. In addition, a "fast adaptive feedback recurrent neural network (FARNN)" for pre-virtual CPU allocation and a modified pigeon-inspired optimization (MPIO) approach for task clustering were developed. A "deep convolutional neural network (DCNN)" is used to power the task load monitoring system, which is designed to provide dynamic priority-based scheduling.

A "self-adapting task scheduling algorithm" in short titled as ADATSA utilising learning automata was suggested by Lili Zhu et al., [17] to solve environmental dynamics such complicated dependence relationships and quick iterations of container scheduling. They developed a useful reward-penalty method for scheduling activities in conjunction with the present environment's running tasks as well as idle resource states.

A "self-adaptive deep reinforcement learning-based scheduler (RLSched)" that automatically captures the resource utilisation dynamics in the data centre has been developed by T. L. Botran et al., [18]. A decentralised actor-critic multi-agent architecture, which allows for parallel execution and quicker convergence, served as the foundation for this scheduler. RLSched is based on an improved network architecture including action shaping that filters out improper activities and keeps the agent from adopting a poor policy.

However, most of the contemporary approaches endeavoured to maximize the performance of the container utilization, they are centric to resource management problems but not considering the contextual dynamics raised during the containers scheduling. Hence, it is obvious to encourage the present research towards container scheduling with regard to contextual dynamics such as roundtrip time, energy consumed and available, as well as computational-load processing time.

To the best of our knowledge, the suggested technique QASC is the first effort to investigate the efficacy of contextual dynamics of the containers using a simple-moving-average that is extremely scalable. QASC is centred to the contextual dynamics of the containers, as opposed to other ideas, since it focuses on optimising the usage of the containers serviced and hence increases predicted performance.

**2. Methods and Materials.** In the context of scheduling containers to delay-sensitive tasks, the Quality Aware Scheduling of Containers (QASC) model assumes a significant role in ensuring task completion within the required time frames while optimizing resource usage. QASC is a method for streamlining the process of allocating jobs to available containers in a computational system as shown in figure 2.1. The main objective of QASC is to increase scheduling efficiency while taking into account a variety of variables that may have an impact on system performance as a whole. This system accepts a set of containers as input, along with their corresponding logs of makespans (the total amount of time required to complete a schedule) and specific metrics for each container, such as I/O-intensive workload, startup time, failure rate, and inter-container dependencies. QASC establishes initial, minimum, maximum, and end values in a normalized form for each container and each metric. In order to track trends over time, it then computes the simple moving average of each of these progression measures. Further exploring these moving averages' patterns, QASC uses them to determine a metric coefficient for each measure. This coefficient offers a thorough evaluation of a container's performance by combining the mean and root mean square distance of progression measures. Additionally, the system determines a quality coefficient for every container that provides an estimate of the metric-coefficients threshold. This threshold essentially assesses the effectiveness of scheduling while taking into account the unique demands of each container and its characteristics. The figure 2.1 outlines the process of the Quality Aware Scheduling of Containers (QASC) system. Here's a detailed explanation:

**User Input:** The user provides the QASC system with a set of containers (CT) and their corresponding logs of make-spans, as well as specific metrics for each container. These metrics include:

I/O Intensive Workload (ioiw)

Container Startup Time (cst)

Container Failure Rate (cfr)

Inter-container Dependencies (icd) **Normalized Progression Measure Calculation:** For each container and its respective metrics, the system calculates normalized progression measures ( $x_{norm}_{ct}$ ).

**Simple Moving Average (SMA) Calculation:** The system computes the simple moving average of each of the progression measures for every metric.

**Progress Measurement Patterns:** It then discovers patterns in the progression measurement.

**Metric Coefficient Calculation:** After identifying patterns, the system calculates the Metric Coefficient for each metric. This involves computing the mean () and the root mean square distance (RMSD) of progression

measures. The metric coefficient (mc) for each metric is computed as the sum of its mean and RMSD.

**Quality Coefficient Calculation:** The system calculates the Quality Coefficient for each container.

**Metric-Coefficients Threshold Estimation:** The system estimates the metric-coefficients threshold ( $qc_{ct}$ ) for each container as a function of the metric coefficients and specific metrics.

**Residual Energy Estimation:** For each container, the system estimates the residual energy ( $RE_{ct}$ ), which is the energy left after accounting for the energy consumption ( $E_{ct}$ ) and the product of energy coefficient ( $EC_w$ ) and the container-specific weight ( $w_{ct}$ ).

**Container Ranking:** The containers are then ranked in descending order based on their metric coefficient thresholds ( $qc_{ct}$ ).

**Container Selection:** Based on the ranking, the system selects the optimal container for scheduling.

**Optimal Scheduling:** Finally, the optimal scheduling of containers is performed, completing the QASC process.

The capability of QASC to calculate each container's residual energy, which represents the container's ability to fulfill the request, is another important feature. Additionally, the system figures out energy consumption, which is the difference between the starting point and the anticipated residual value.

The best container for scheduling is then chosen after containers are ranked according to their metric coefficient thresholds. Thus, the QASC approach considers a number of variables to optimize container scheduling, with the goal of enhancing system performance as a whole.

For analysis and predictive scheduling, the QASC model computes Simple Moving Averages (SMAs) based on beginning value, maximum value, and minimum representational form. SMA allows for the following: beginning times, failure rates, determining basic I/O burden, smoothing volatility, and dependency on other containers. Understanding how containers function over long distances and minimizing the effects due to transient anomalies make smoothing important. These SMA values may be used by QASC to more precisely predict the susceptibility of containers. As a result, job distribution might be based on the functional profile of each container from inception to completion, baseline, and peak. So that no container is under scheduled due to a conservative policy or overloaded when demand increases, these predictive insights are important in resource allocation. This significantly raises the reliability and efficiency of cloud computing.

Here are the QASC factors in the context of delay-sensitive tasks:

- **I/O Intensive Workload:** The I/O operations handled by a container during the average makespan time for delay-sensitive tasks are considered. Containers with higher I/O workloads might experience higher latency, which could potentially affect delay-sensitive tasks.
- **Container Startup Time:** The startup time of a container is a critical factor when dealing with delay-sensitive tasks. A container with a shorter startup time can begin processing tasks faster, reducing the risk of delays.
- **Container Failure Rate:** A container with a higher failure rate may disrupt the timely completion of tasks. This is especially crucial for delay-sensitive tasks, where any failure can lead to unacceptable delays.
- **Inter-container Dependencies:** The dependencies between containers can also impact the scheduling of delay-sensitive tasks. A container with fewer dependencies is less likely to be delayed due to issues with other containers, which makes it more suitable for delay-sensitive tasks.

The Quality Score ( $QS$ ) calculated based on these metrics guides the QASC model in allocating tasks to containers. Containers with higher  $QS$ , indicating lower I/O workload, faster startup time, lower failure rate, and fewer dependencies, are prioritized for delay-sensitive tasks to ensure timely completion and optimize the usage of resources.

**2.1. Quality Aware Scheduling of Containers (QASC). Input:** Set of containers  $CT$ , their corresponding log of makespans, container specific metrics: (1). I/O Intensive Workload ( $ioiw$ ), (2). Container Startup Time ( $cst$ ), (3). Container Failure Rate ( $cfr$ ), and (4). Inter-container Dependencies ( $icd$ ). **Output:** Optimal scheduling of containers. **Initialization:** For each container  $ct \in CT$ , For each makespan  $m \in log$  that successfully completed, For each metric  $i \in ioiw, cst, cfr, icd$ , Determine the initial ( $o$ ), minimum ( $l$ ), maximum ( $h$ ), and end ( $e$ ) value of the metric in normalized form as follows. For each metric  $i \in ioiw, cst, cfr, icd$  and for each container  $ct \in CT$ , Calculate the normalized progression measure  $x_{norm}$  for initial ( $o$ ), min ( $l$ ), max

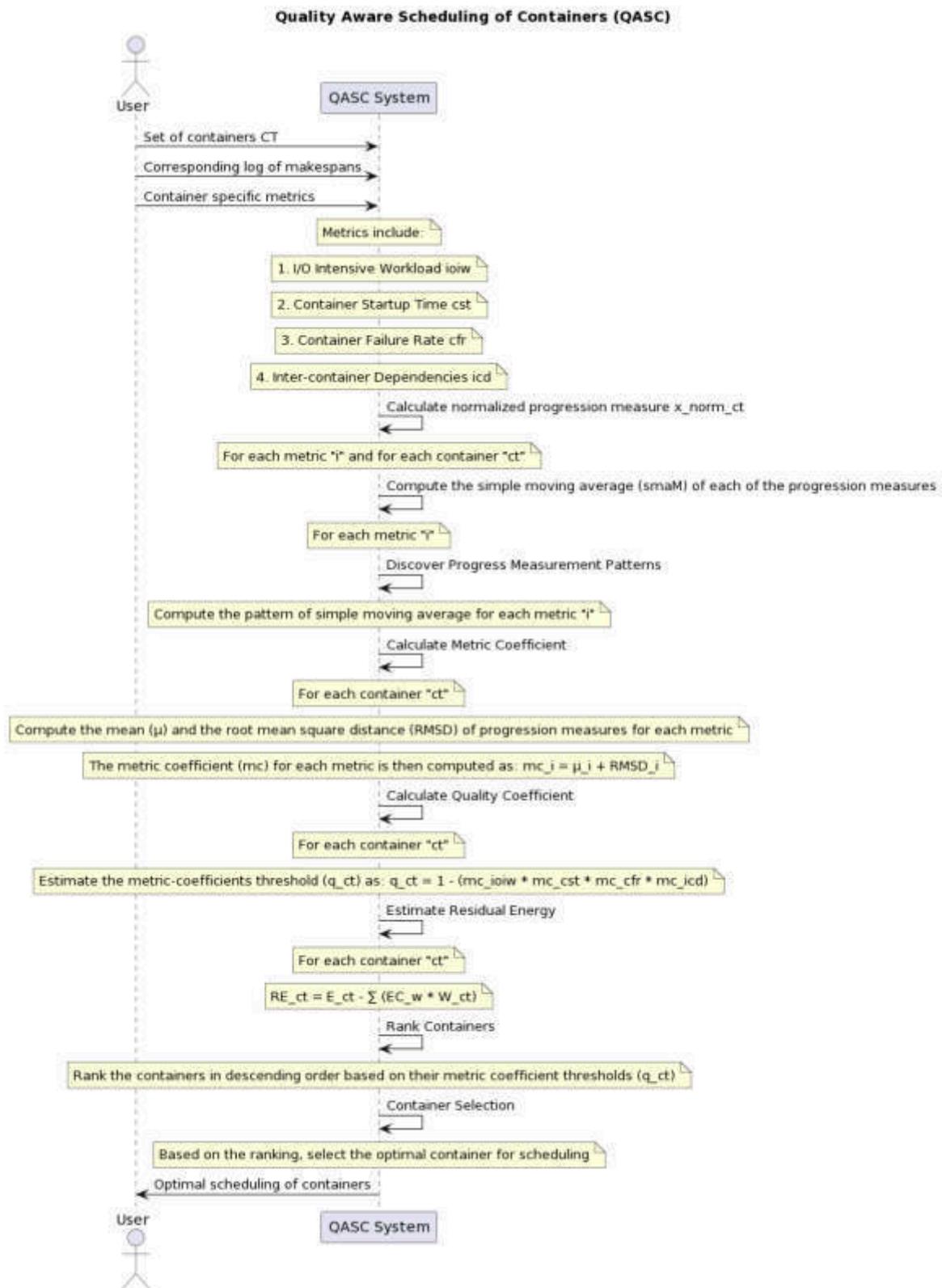


Fig. 2.1: The flow diagram representation of the QASC

( $h$ ), and end ( $e$ ) values using the formula: Eq 2.1

$$x_{(norm_{ct})} = 1 - (1/x_{ct}) \quad (2.1)$$

where  $x_{ct}$  is the original progression measure for the metric  $i$  in the container  $ct$ , and  $x_{norm_{ct}}$  is the normalized progression measure. Compute the simple moving average of each of the progression measures ( $o$ ,  $l$ ,  $h$ ,  $e$ ), for each metric  $i$  as  $smaM_i$  follows: For a set of metrics  $M = m_1, m_2, \dots, m_n$  and a simple moving average configuration  $smac$ , the simple moving average of the  $i^{th}$  metric's initial values, maximum values, minimum values, and end values can be represented by four equations:"

- Simple moving average of initial values ( $sma_i(o)$ ):  $sma_i(o) = (1/smac) * \sum_{j=i}^{i+smac} m_j(o)$
- Simple moving average of maximum values ( $sma_i(h)$ ):  $sma_i(h) = (1/smac) * \sum_{j=i}^{i+smac} m_j(h)$
- Simple moving average of minimum values ( $sma_i(l)$ ):  $sma_i(l) = (1/smac) * \sum_{j=i}^{i+smac} m_j(l)$
- Simple moving average of end values ( $sma_i(e)$ ):  $sma_i(e) = (1/smac) * \sum_{j=i}^{i+smac} m_j(e)$

where:

- $m_j(o)$ ,  $m_j(h)$ ,  $m_j(l)$ ,  $m_j(e)$  represent the initial value, maximum value, minimum value, and end value of the  $j^{th}$  metric respectively.
- The summation  $\sum_{j=i}^{i+smac}$  indicates that the sum is taken from the  $i^{th}$  metric to the  $(i + smac)^{th}$  metric.

Finally, the set of these simple moving averages  $smaM$  for all metrics  $i$  in  $M$  is given by:  $smaM = sma_i(o)$ ,  $sma_i(h)$ ,  $sma_i(l)$ ,  $sma_i(e)$  for all  $i$  in  $M$

Discover Progress Measurement Patterns:

For each metric  $i \in ioiw, cst, cfr, icd$ ,

Compute the pattern of simple moving average for each progression measure as:

- $psmaM_i(o) = smaM_i(o)$  if  $smaM_i(o)$  belongs to the first makespan
- Else,
- $psmaM_i(o) = (psmaM_{(i-1)}(o) + psmaM_{(i-1)}(e))/2$
- $psmaM(e) = (smaM(o) + smaM(h) + smaM(l) + smaM(e))/4$
- $psmaM(l) = \min(smaM(l), psmaM(o), psmaM(e))$
- $psmaM(h) = \max(smaM(h), psmaM(o), psmaM(e))$

**Calculate Metric Coefficient:** For each container  $ct \in CT$ , Compute the mean ( $\mu$ ) and the root mean square distance (RMSD) of progression measures for each metric as

$$mu_i = (psmaM_i(o) + psmaM_i(h) + psmaM_i(l) + psmaM_i(e))/4 \quad (2.2)$$

$$RMSD_i = \sqrt{((psmaM_i(o) - \mu_i)^2 + (psmaM_i(h) - \mu_i)^2 + (psmaM_i(l) - \mu_i)^2 + (psmaM_i(e) - \mu_i)^2)/4} \quad (2.3)$$

The metric coefficient ( $mc$ ) for each metric is then computed as

$$mc_i = \mu_i + RMSD_i \quad (2.4)$$

**Calculate Quality Coefficient:** For each container  $ct \in CT$ , Estimate the metric-coefficients threshold ( $q_{ct}$ ) as

$$q_{ct} = 1 - (mc_{ioiw} * mc_{cst} * mc_{cfr} * mc_{icd}) \quad (2.5)$$

**Estimate Residual Energy:** For each container  $ct \in CT$ , Let's denote:

$E_{ct}$  as the current energy level of container  $ct$   $W_{ct}$  as the workload of container  $ct$   $EC_w$  as the estimated energy consumption of workload  $w$  Then, the Residual Energy (RE) of a container  $ct$  can be calculated as follows

$$RE_{ct} = E_{ct} - \sum (EC_w * W_{ct}) \quad (2.6)$$

Estimate the residual energy ( $re_{ct}$ ) that indicates the container's scope to complete the request. Also, estimate the energy consumption ( $ec_{ct}$ ) which is the unconditional disparity between the initial-value and the predicted residual-value.

Table 3.1: Record of Progress Measures

Container	Make span	Metric	Initial	Maximum	Minimal	Residual
-----------	-----------	--------	---------	---------	---------	----------

Table 3.2: Mean Make-Span Ratio of the QASC, DSTS, ADATSA, and RLSched

No. of Containers	QASC	DSTS	ADATSA	RLSched
10	31	18	15	12
15	33	20	17	14
20	36	20	17	14
25	39	24	21	18
30	41	32	29	26
35	44	33	30	27
40	44	33	30	27
45	47	33	30	27
50	49	34	31	28

**Rank Containers:** Rank the containers in descending order based on their metric coefficient thresholds ( $q_{ct}$ ).

**Container Selection:** Based on the ranking, select the optimal container for scheduling.

The quality of the scheduling is determined by considering factors such as the container's I/O workload, startup time, failure rate, inter-container dependencies, and energy consumption. The selection of the optimal container is done based on the calculated metric and quality coefficients, and the anticipated residual energy [19]."

**3. Experimental study.** The goal of the experimental investigation was to expand the efficiency of the suggested scheduling model in order to schedule containers to requests. The experiments were conducted on a dataset [20], which represents the metrics of each make-span in series observed from each container. The dataset shows a collection of records with the following table 3.1 structure. Cloud-sim based simulation has been used to create the dataset mentioned above. The dataset was created by taking into account all 1,200 internet protocol-capable users. Through the scheduling gateway, this Internet of Things network is connected to the resource provider. Furthermore, a cloud computing network with 50 containers is connected to these resource providers. Each container's most recent 20 make-spans are used to compile the metric values. The dataset generated has a total of 20,000 records included in it. The average make-span-ratio, roundtrip-time against a varied count of containers, as well as adaptable transactional requests load are the parameters taken into account for performance study. Energy consumption ratio versus fluctuating load, another key metric, has also been evaluated.

The comparison of metric values acquired from the QASC with the comparable results from the existing models DSTS [16], ADATSA [17], as well as RLSched [18], denoting the efficiency of the QASC.

The amount of time needed to do the activity is indicated by the metric "make-span ratio". The make-span ratio is predicted to be proportional to the count of containers in Table 3.2, Figure 3.1. According to QASC scheduling model, the mean make-span ratio of  $21.7 \pm 5.1$  is being noted. In contrast to the mean make-span ratios of the DSTS ( $24.4 \pm 3.4$ ), ADATSA ( $27.5 \pm 3.4$ ), and RLSched model ( $40.4 \pm 0.4$ ), the mean make-span ratio of the QASC is noticeably high and has a little variance.

For QASC, DSTS, ADATSA, and RLSched, another quality metric stated as roundtrip time is being taken into account and compared (see Table 3.3, Figure 3.2). Figure 3.2 statistics, which show that the average roundtrip times observed from the models QASC, DSTS, ADATSA, and RLSched are respectively  $20.4 \pm 6.6$ ,  $36.5 \pm 6.4$ ,  $40.5 \pm 10.4$ , and  $44.5 \pm 14.4$ , indicate that the model QASC outperforms the existing models DSTS, ADATSA, as well as RLSched with shorter roundtrip times. In Table 3.4, Figure 3.3, the roundtrip time under varied load is predicted. Comparing the performance of the model QASC to that of the existing methods DSTS,

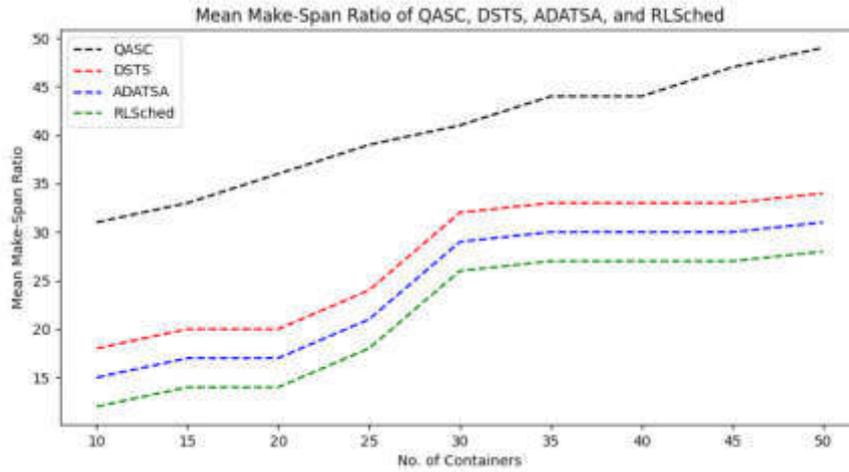


Fig. 3.1: The mean make-span ratio of the QASC, DSTS, ADATSA, and RLSched

Table 3.3: Between a Fixed Load and an Adaptable Number of Containers, the Mean Roundtrip Time

No. of Containers	QASC	DSTS	ADATSA	RLSched
10	24	45	49	53
15	29	43	47	51
20	26	43	47	51
25	25	40	44	48
30	22	36	40	44
35	20	34	38	42
40	17	32	36	40
45	13	31	35	39
50	7	25	29	33

ADATSA, as well as RLSched, it is clear that it performs better. The mean difference among the roundtrip times of the QASC ( $18.9 \pm 8.1$ ) and the DSTS ( $34.8 \pm 7.95$ ), the ADATSA ( $39.8 \pm 12.9$ ), and the RLSched model ( $44.8 \pm 17.9$ ) has been noted.

Table 3.5, Figure 3.4 shows the data of the important objective energy use. Comparisons are made between the proposed QASC and the current DSTS, ADATSA, as well as RLSched models' average energy usage (measured in joules) by each of the make-spans. The outcome demonstrates that the suggested model QASC outperforms the current models DBSA, which consumes an average of  $20 \pm 11$  joules per make-span, ADATSA, which consumes an average of  $24 \pm 15.22$  joules per make-span, and RLSched, which consumes an average of  $28 \pm 19.22$  joules per make-span. In comparison to the existing methods DSTS, ADATSA, and RLSched, the QASC uses less energy on average.

The improvement in the performance can be understood as a result of QASC's holistic approach, which takes into account not only the current state of the containers but also their historical performance data. This enables the QASC system to make more informed and precise scheduling decisions, leading to the observed enhancements in operational efficiency and task handling capacity.

**4. Conclusion.** The QASC has been put through a great deal of testing in cloud computing operations and is a substantial upgrade over existing models such as DSTS, ADATSA, and RLSched. This is demonstrated by the empirical findings of extensive experimental testing. By taking into account make-span, round-trip

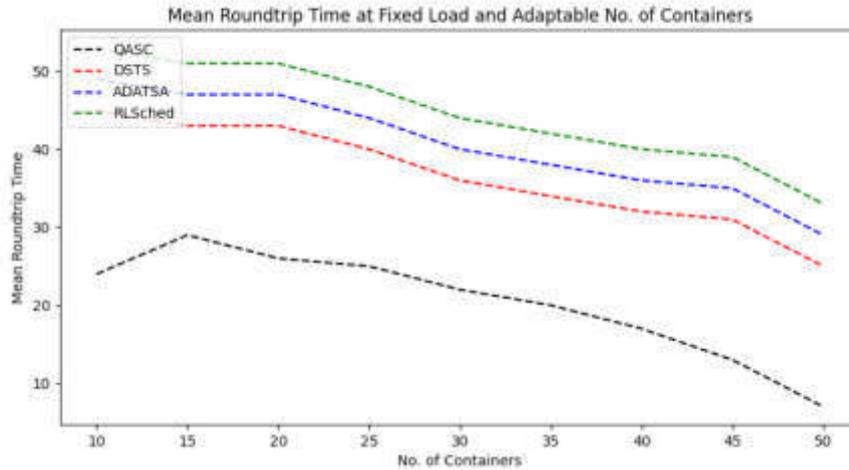


Fig. 3.2: The mean roundtrip time of QASC, DSTS, ADATSA, and RLSched observed from adaptable count of containers

Table 3.4: Mean Roundtrip Time Observed at Fixed Count of Containers vs. Adaptable Count of Requests

Adaptable Count of Requests in KBPS	QASC	DSTS	ADATSA	RLSched
10	7	24	29	34
15	10	27	32	37
20	13	30	35	40
25	18	31	36	41
30	18	31	36	41
35	19	36	41	46
40	22	42	47	52
45	29	44	49	54
50	34	49	54	59

time, energy consumption, and container properties, QASC's scheduling method enhances performance. The translation look-ahead approach lowers make-span ratios and roundtrip times when container counts change. Energy efficiency, a key component of green cloud computing, is where QASC beats existing cloud computing models. According to the energy consumption data, QASC consumes less energy for different request loads per make-span. Such efficiency is needed for large-scale cloud computing to be sustainable both financially and environmentally. Due to its efficiency in resource allocation and high quality, QASC is helpful for cloud computing. This demonstrates how balancing resource consumption and performance may greatly increase cloud efficiency. Lastly, QASC promises a bright future by enhancing cloud computing resource management performance and energy consumption. It may take the lead in cloud computing technologies in the future due to its outstanding performance and better results than those of existing models.

#### REFERENCES

- [1] I. AHMAD, ET AL., *Container scheduling techniques: A survey and assessment*, Journal of King Saud University-Computer and Information Sciences, 34.7 (2022), pp. 3934-3947.
- [2] K. SENJAB, S. ABBAS, AND N. AHMED, *A survey of Container scheduling algorithms*, Journal of Cloud Computing, 12.1 (2023), pp. 1-26.
- [3] S. SINGH, AND I. CHANA, *A survey on resource scheduling in cloud computing: Issues and challenges*, Journal of grid

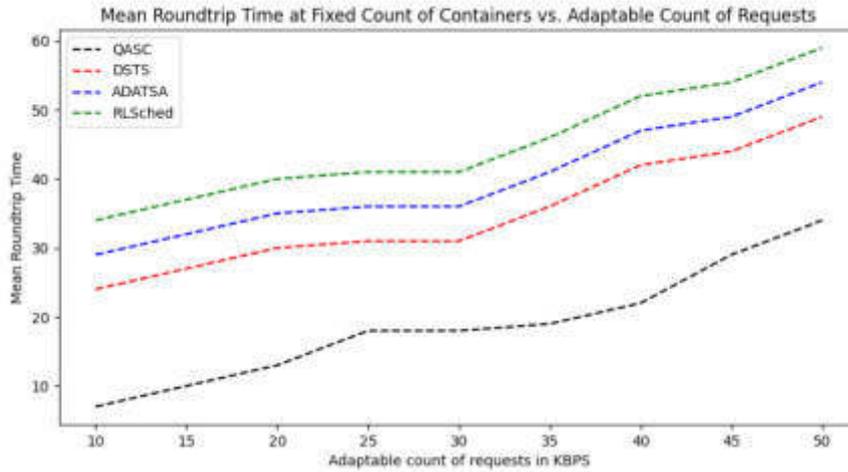


Fig. 3.3: Roundtrip time observed from the QASC, DSTS, ADATSA, and RLSched against adaptable load of requests

Table 3.5: Energy Usage (Measured in Joules) by Each of the Make-Spans at Adaptable Load of Requests

Adaptable Load of Requests in KBPS	QASC	DSTS	ADATSA	RLSched
10	1	3	7	11
15	2	8	12	16
20	7	11	15	19
25	10	17	21	25
30	14	19	23	27
35	17	22	26	30
40	19	27	31	35
45	21	34	38	42
50	25	39	43	47

computing, 14 (2016), pp. 217-264.

- [4] Kubernetes, *Google Kubernetes*, URL: <https://kubernetes.io/docs/concepts/schedulingeviction/kube-scheduler/> (visited on 01/2021).
- [5] Docker, *Docker Swarm*, URL: <https://docs.docker.com/engine/swarm/how-swarm-modeworks/services/> (visited on 01/2021).
- [6] S. LU, M. NI, AND H. ZHANG, *The optimization of scheduling strategy based on the Docker swarm cluster*, *Information Technology*, 40.7 (2016), pp. 147–151.
- [7] D. ZHANG, ET AL., *Container oriented job scheduling using linear programming model*, In: 2017 3rd International Conference on Information Management (ICIM), IEEE, 2017, pp. 174–180.
- [8] C. KAEWKASI, AND K. CHUENMUNEEWONG, *Improvement of container scheduling for docker using ant colony optimization*, In: 2017 9th international conference on knowledge and smart technology (KST), IEEE, 2017, pp. 254–259.
- [9] R. ZHANG, ET AL., *A genetic algorithm-based energy-efficient container placement strategy in CaaS*, *IEEE Access*, 7 (2019), pp. 121360–121373.
- [10] A. MSIEDDI, ET AL., *Joint container placement and task provisioning in dynamic fog computing*, *IEEE Internet of Things Journal*, 6.6 (2019), pp. 10028–10040.
- [11] Z. WANG, ET AL., *Automated cloud provisioning on aws using deep reinforcement learning*, arXiv preprint arXiv:1709.04305, 2017.
- [12] B. DU, C. WU, AND Z. HUANG, *Learning resource allocation and pricing for cloud profit maximization*, In: Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 33, 2019, pp. 7570–7577.
- [13] S. ZHANG, ET AL., *A-SARSA: A Predictive Container Auto-Scaling Algorithm Based on Reinforcement Learning*, 2020 IEEE International Conference on Web Services (ICWS), IEEE, 2020, pp. 489–497.
- [14] Q. ZHANG, ET AL., *A double deep Q-learning model for energy-efficient edge scheduling*, *IEEE Transactions on Services*

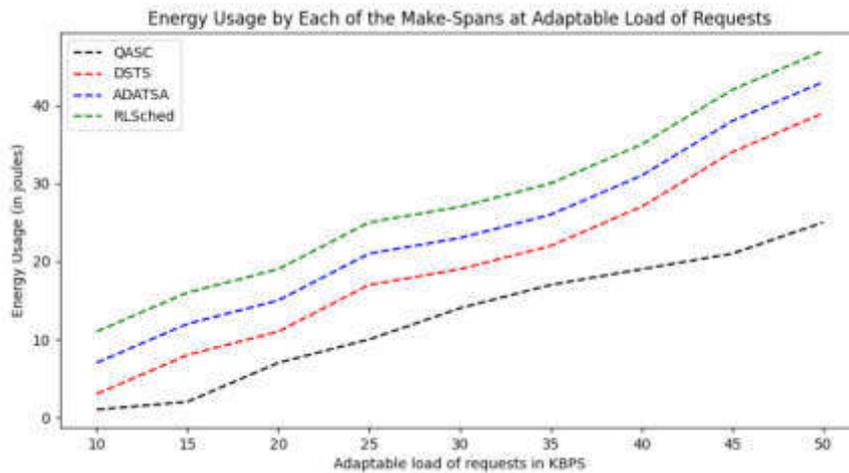


Fig. 3.4: The energy used (measured in joules) by each of the make-spans noted from the QASC, DSTS, ADATSA, and RLSched techniques

- Computing, 12.5 (2018), pp. 739–749.
- [15] W. FUNIKA, J. KOPEREK, AND P. KITOWSKI, *Automatic management of cloud applications with use of Proximal Policy Optimization*, International Conference on Computational Science, Springer, 2020, pp. 73–87.
- [16] S. MUNISWAMY, AND R. VIGNESH, *DSTS: A hybrid optimal and deep learning for dynamic scalable task scheduling on container cloud environment*, Journal of Cloud Computing, 11.1 (2022), pp. 1–19.
- [17] L. ZHU, ET AL., *A Self-Adapting Task Scheduling Algorithm for Container Cloud Using Learning Automata*, IEEE Access, 9 (2021), pp. 81236–81252.
- [18] T. LORIDO-BOTRAN, AND M. K. BHATTI, *Adaptive container scheduling in cloud data centers: a deep reinforcement learning approach*, International Conference on Advanced Information Networking and Applications, Springer, Cham, 2021.
- [19] A. GELLERT, A. FLOREA, U. FIORE, F. PALMIERI, AND P. ZANETTI, *A study on forecasting electricity production and consumption in smart cities and factories*, International Journal of Information Management, 49 (2019), pp. 546–556.
- [20] Kaggle, *Datasets*, Available at: <https://www.kaggle.com/datasets>.

*Edited by:* Anil Kumar Budati

*Special issue on:* Soft Computing and Artificial Intelligence for wire/wireless Human-Machine Interface

*Received:* Oct 4, 2023

*Accepted:* Jan 18, 2024



## ENHANCED FEATURE OPTIMIZATION FOR MULTICLASS INTRUSION DETECTION IN IOT FOG COMPUTING ENVIRONMENTS

SUDARSHAN S. SONAWANE\*

**Abstract.** To overcome the shortcomings of traditional security measures in IoT fog computing, The Multiclass Intrusion Detection (MCID) model is put forward. The model's goal is to improve intrusion detection by identifying and classifying different attack types. The behavioral, temporal and anomaly features are fused through SVM-BFE for obtaining the best possible selection of high worth features. Finally we use a Random Forest algorithm to robustly classify them. There is also its adaptability to the ever-changing security demands of fog computing. MCID's ability to improve the cloud security of fog computing is shown by a 4-fold cross validation, which returns performances including precision rates up to 99.43 %, recall about 95 % and F-measures at as much as 97.17 %. Moreover there are specificity rate totals coming in over this whole range that hit close or right

**Key words:** Internet of Things (IoT); Intrusion Detection System; UNSW-NB15; Fog Computing; SVM-BFE

**1. Introduction.** The Internet of Things (IoT) has grown quickly, integrating a wide range of devices and facilitating easy data exchange and communication [1]. This evolution calls for the need for an intermediary computing layer between cloud computing and IoT devices [2]. By processing data closer to its source and increasing efficiency, fog computing was able to solve this issue. Fog computing crosses over devices with disparate operating systems, which presents a security risk. Fog computing raises the stakes for security, especially for intrusion detection. In decentralized environments, traditional detection systems that are built for centralized computing architectures perform poorly [3]. The use of IoT in healthcare, transportation, and defense makes security issues more important to handle.

This research aims to enhance intrusion detection, specifically for fog computing, in light of these challenges. The MCID model takes this into account. Behavioral, temporal, and anomaly detections are all part of the MCID model. The model optimizes feature selection using SVM-BFE optimization to enhance detection. To evaluate the applicability of the MCID model, we conducted extensive testing against established benchmarks. The NUSW-NB15 dataset offers a strong basis for evaluation. For testing, we also employed the Random Forest algorithm. MCID's performance is evaluated using this method, which contrasts it with other models. The MCID model aims to enhance fog computing security. Ensuring a robust intrusion detection framework is our goal in order to increase IoT safety and reliability.

This article follows a logical structure. Following a discussion of fog computing's challenges, we evaluate the MCID model's performance and implications. In order to explain Fog Computing security challenges and solutions, this organized approach tries to highlight the critical role that the MCID model plays in addressing them.

**2. Related Work.** We face both tremendous security challenges and exciting opportunities as fog computing and the Internet of Things grow quickly. Risks increase as more devices connect and systems get more complicated. Recent research on these risks is examined in the "Related Research" section. New threat mitigation methods have been proposed, the flaws of older security systems in fog environments have been investigated, and their effectiveness has been put to the test. Everything from basic, efficient security to sophisticated deep learning protection is covered in this section.

Belal Sudqi Khater et al. [4] focused on addressing resource limitations in traditional intrusion detection systems when applied in fog computing environments. Fog computing extends cloud capabilities to network edges using diverse hardware devices, posing challenges for resource-intensive intrusion detection systems. The article's objective is to propose a lightweight intrusion detection system capable of running effectively on

---

\*Department of Computer Engineering, R. C. Patel Institute of Technology, Shirpur, India ([sudars2000@gmail.com](mailto:sudars2000@gmail.com))

resource-constrained fog devices while maintaining high detection rates with minimal computational overhead. To achieve this, the authors introduce a novel perceptron-based intrusion detection system that utilizes vector space representation and a multilayer perceptron model. Experimental results reveal a remarkable achievement, with the system delivering a 94% Accuracy, 95% Recall, and 92% F1-Measure in one dataset and 74% Accuracy, 74% Recall, and 74% F1-Measure in another. Moreover, the system operates efficiently on a Raspberry Pi 3 fog device, with low energy consumption, demonstrating its suitability for fog computing environments.

KISHWAR SADAF et al. [5] proposed a novel approach using deep learning techniques, Autoencoder and Isolation Forest. The objective is to differentiate between normal and attack packets in real-time with high accuracy, enhancing the security of fog devices. The article introduces the Auto-IF method, combining Autoencoder and Isolation Forest, to process incoming network packets efficiently and detect malicious attacks in real-time while minimizing false positives. The authors extensively evaluate the method, utilizing the NSL-KDD dataset, and report superior performance compared to other intrusion detection approaches. Notably, the Auto-IF method effectively reduces false positives and negatives, particularly when the dataset's contamination percentage is around 10-12%. This contribution offers an efficient solution to bolster the security of fog devices and ensure reliable service delivery.

Prabhat Kumar et al. [6] confronted the escalating security concerns surrounding Internet of Things (IoT) networks. The exponential growth of connected IoT devices brings with it an increased risk of attacks and compromise. These devices, often resource-constrained, are susceptible to security breaches that can lead to unauthorized data access, physical infrastructure damage, or their enlistment in botnets for further attacks. Traditional security measures prove inadequate for safeguarding IoT networks. Consequently, the article introduces a distributed ensemble design-based intrusion detection system, leveraging machine learning and fog computing to enhance intrusion detection's accuracy and efficiency in IoT networks. The aim is to protect these networks from diverse attacks and bolster their security.

K. Kalaivani et al. [7] tackled the security challenges prevalent in cloud and fog computing environments, which are vulnerable to various types of attacks such as DDoS, Worm, Probe, R2L, and U2R. These attacks can disrupt network resources and consume bandwidth, posing significant threats. Conventional methods often fall short in accurately detecting and countering these attacks. The article presents a novel solution—a hybrid deep learning intrusion detection model named ICNN-FCID. This model combines Intrusion Classification Convolutional Neural Network (ICNN) with Long Short-Term Memory (LSTM) networks and is specifically designed for fog computing environments. ICNN-FCID demonstrates remarkable effectiveness, achieving an accuracy of about 96.5% and effectively classifying various attack types. By proposing this model, the article contributes to enhancing security in cloud and fog computing environments.

Mohammed Anbar et al. [8] addressed the imperative need for effective intrusion detection systems (IDSs) in Fog computing. Fog computing, an extension of cloud computing, shifts computational resources closer to network edges, reducing communication costs and boosting system performance. However, its distributed nature poses new security challenges, necessitating robust IDSs. This article conducts a comprehensive review of recent IDS research in Fog computing, categorizing various approaches and highlighting their challenges and limitations. The review provides a new taxonomy of IDSs, outlines shortcomings such as scalability and resource constraints, and offers potential solutions and future research directions. In sum, this article contributes to an improved understanding of IDSs in the Fog computing landscape, guiding further research in this vital domain.

Hassan Adegbola Afolabi et al. [9] addressed the pressing security and privacy concerns surrounding IoT devices connected via Fog computing. These devices, due to limited resources and inadequate security, are susceptible to various threats like DDoS attacks and brute-force attacks. The article proposes an Intrusion Detection System (IDS) model utilizing Back Propagation Deep Neural Network (BP-DNN) to safeguard the Fog-IoT platform. The model aims to enhance accuracy and detection rates, mitigating security breaches in IoT devices within Fog computing.

M. Ramkumar et al. [10], proposed a Support Vector Machine (SVM)-based Intrusion Detection System (IDS) tailored for the unique challenges of fog computing. The SVM algorithm is employed for its efficiency, and the IDS focuses on characteristics like mean, limit, and median systems to detect intrusions efficiently. Offering a lightweight, efficient IDS solution for fog computing with a low false-positive rate and an impressive average detection rate of 98.6% is the contribution.

An extensive survey examining intrusion detection systems for fog and cloud computing is carried out by Victor Chang et al. [11]. The article seeks to shed light on various aspects of security policy deployment in these kinds of environments. Software-as-a-Service (SaaS) and intrusion detection are highlighted as methods for tracking and evaluating network traffic, and strategies and policies for lowering intrusion detection are also suggested. Despite lacking its own experimental results or statistics, the article is still a useful resource for understanding security policy deployment in fog and cloud computing environments.

Within fog computing, Wu et al. [12] presented a novel Intrusion Detection System (IDS) designed specifically for Internet of Things networks. With Convolutional Neural Network (CNN) as the first classifier and Random Forest as the second, this IDS uses a distributed ensemble design. By using this method, the efficiency and accuracy of identifying malicious attacks on fog nodes will be improved. The suggested model considerably improves detection performance by overcoming the drawbacks of conventional IDS, attaining an astounding 94.43% overall accuracy while effectively allocating the processing load.

The cybersecurity issues that fog computing (FC) and edge computing (EC) environments face because of their dynamic nature and variety of network devices were addressed by Omar A. Alzubi et al. [13]. Traditional intrusion detection systems (IDS) often fall short in these settings. To address this issue, the article proposes an optimized machine learning-based IDS called Effective Seeker Optimization with Machine Learning-Enabled Intrusion Detection System (ESOML-IDS). This system leverages feature selection and parameter optimization through effective seeker optimization and machine learning. Comparative experiments demonstrate the ESOML-IDS's remarkable effectiveness, achieving an accuracy of 99.5% and outperforming other IDS models in FC and EC environments.

Junaid Sajid et al. [14] addressed the susceptibility of UAVs to cyberattacks in smart farming, proposing a fog computing-based framework to bolster security. By combining UAVs and IoT sensors for data collection, a fog computing architecture for data transmission, and an intrusion detection system that employs machine learning, the system identifies compromised UAVs. The framework not only enhances data collection security but also introduces a coin-based recharge system to reward benign UAV behavior while efficiently managing resources.

Kalaivani Kaliyaperumal et al. [15] confronted the vulnerability of fog computing systems to a range of attacks and offer an efficient solution. Their proposed model combines deep learning techniques, such as CNN and IDS-AlexNet, with Random Forest for attack detection. Leveraging feature selection, this model significantly improves attack detection accuracy, achieving an impressive 97.5% accuracy on the UNSW-NB15 dataset. This contribution enhances fog computing security and represents a substantial step toward safeguarding these systems from various threats.

Doaa Mohamed et al. [16] tackled the increasing vulnerability of IoT networks to cyber threats by proposing EHIDS, a hybrid intrusion detection system combining fog and cloud computing. EHIDS aims to identify abnormal behavior in IoT devices and networks to prevent system intrusions. The article details the architecture and implementation of EHIDS and conducts comprehensive experiments to showcase its superior performance compared to other approaches. EHIDS reduces execution time significantly and demonstrates effectiveness in enhancing IoT security.

Cristiano Antonio de Souza et al. [17] provided a comprehensive overview of intrusion detection in IoT and Fog Computing. Their article discusses the challenges, principles of Machine Learning techniques, and state-of-the-art approaches in intrusion detection. While the authors mention experiments with the IoTID20 dataset, they primarily focus on discussing the principles and challenges of intrusion detection, guiding researchers in this field with a rich source of information.

Guosheng Zhao et al. [18] addressed security risks in cloud-fog hybrid computing for IoT by proposing a lightweight intrusion detection model based on ConvNeXt-SF. Their model significantly improves efficiency and performance while maintaining detection and learning capabilities. Detailed experimental results show its superiority, reducing training and prediction times while enhancing accuracy and addressing information security risks effectively. This contribution offers an innovative solution for improving IoT security in cloud-fog hybrid computing environments.

According to research in the cited articles, there are serious security risks associated with the Internet of Things' (IoT) rapid expansion and integration across numerous industries. According to these studies,

IoT intrusion detection systems (IDS) and fog computing are being developed to meet their specific security requirements. The impact of fog environments on traditional IDS resource constraints is discussed by Belal Sudqi Khater et al. [4]. While Prabhat Kumar et al. [6] and M. Ramkumar et al. [10] stress the need for sophisticated algorithms and specialized systems to address vulnerabilities brought on by the growth of connected devices, KISHWAR SADAF et al. [5] use deep learning to improve detection rates.

Different approaches are provided in this domain by the Lightweight Intrusion Detection Model (LIDM) [18] and the Combined Ensemble Intrusion Detection Model (CEIDM) [15]. One layered defense against intrusions is provided by the CEIDM's multiple detection methods. Limitations on IoT resources are addressed by LIDM, which guarantees efficient intrusion detection with low computational overhead.

The proposed MCID model is pertinent and required in light of this situation. Despite being effective, many of the current methods only deal with particular problems. Due to the complexity of fog computing, a broad-spectrum IDS is required. Behavioral, temporal, and anomaly-based detections are addressed by the MCID model, which takes a tri-dimensional approach to fog computing's many challenges. To enhance detection, MCID additionally employs SVM-BFE optimization. MCID's robustness is highlighted by its efficient yet comprehensive solution, which stands in contrast to the lightweight and ensemble methods of CEIDM and LIDM.

The reviewed articles take on key security issues in the rapidly expanding IoT and fog computing space. To avoid resource constraints in fog devices, Khater et al. [4] and Afolabi et al. [9] suggest lightweight IDS solutions. Deep learning is also used by Sadaf et al. [5] to improve real-time attack detection using AutoIF and Kalaivani et al. [15], implementing ICNNFCID respectively. This leads to greater accuracy and fewer false positives in the methods. Kumar et al. [6] and Ramkumar et al., advocate for machine learning-based IDS, with consideration towards distributed designs (node autonomous monitoring) to guard against multiple attacks of different types. In addition the researchers focus on efficiency in SVMs based filtration mechanism Anbar et al. [8] and Zhao et al. [18], provide detailed overviews, stressing the importance of an effective policy for security in fog computing systems and cloud communities, respectively. Wu et al. [12] and Alzubi et al. [13], respectively, propose new IDS models utilizing ensemble learning technologies or optimization techniques that have high accuracy and impressive performance records. Sajid et al. [14] proposed a fog computing framework for securing UAVs in smart farming systems. Furthermore, Mohamed et al. [16] and Souza et al. [17] look at hybrid intrusion detection as well as the basics of machine learning in IoT security, demonstrating how hard it is to detect attacks on an IoT network.

Based on the literature reviewed, it is apparent that the MCID framework stands out as an important contribution to the area of IoT fog computing security. With fog computing, intrusion detection is a unique challenge. The research landscape shows that there are various attempts to solve it; these efforts focus on resource efficiency, real-time ability and accuracy. Yet, there still is a need for an integrated solution that solves these individual problems. This is the integrated approach offered by MCID, which integrates state-of-the-art feature optimization and classification techniques. It promises to avoid some of the limitations of existing models while allowing for a robust system that will be more than adequate for use in highly complex fog computing environments.

**3. Methods and Materials.** The proposed model addresses a critical need in the realm of IoT fog computing, where the amalgamation of diverse computing entities and the decentralized nature of fog computing render traditional security measures insufficient. The convergence of myriad devices and the ensuing data explosion necessitate advanced intrusion detection mechanisms that can adeptly identify and categorize varied intrusion types, ensuring the security and reliability of fog computing ecosystems. MCID, with its intricate feature optimization and multiclass classification capabilities, is not merely a response to this need but a pioneering initiative, offering nuanced insights and enhanced detection accuracies in distinguishing between normal and malicious activities. The significance of MCID lies in its ability to harness advanced machine learning techniques, optimizing feature selection using SVM-BFE and employing ensemble learning for refined classification, thereby contributing to the fortification of IoT fog computing environments against escalating and evolving cyber threats. The justification for deploying this model is underscored by the urgent imperative to secure multifaceted and dynamic fog computing landscapes, where conventional security protocols are rendered obsolete, and the demand for sophisticated, adaptive, and robust intrusion detection mechanisms is ever-increasing.

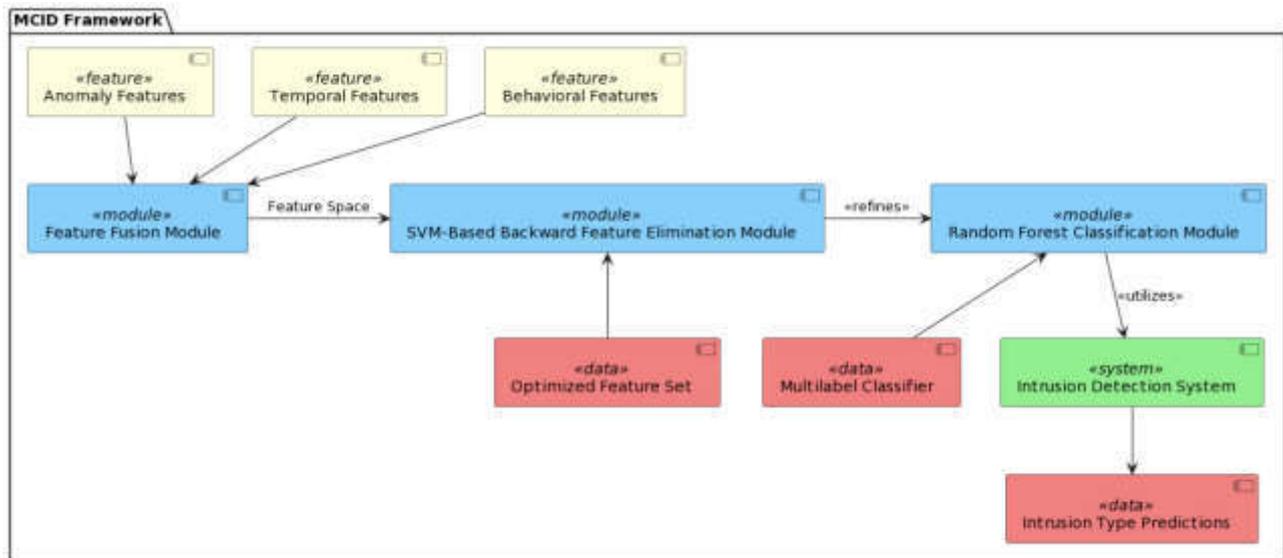


Fig. 3.1: MCID Architecture

In essence, MCID stands as a beacon of innovation, promising enhanced security and resilience in IoT fog computing, paving the way for the safe and seamless integration of fog computing in our interconnected digital world.

The MCID architecture that shown in figure 3.1 is designed to optimize feature selection and enhance intrusion detection capabilities in IoT fog computing environments. This architecture is embedded with a fusion of Behavioral, Temporal, and Anomaly features to construct a multidimensional feature space capable of detailed representation of data patterns and nuances, enabling heightened discernibility between varied intrusion types and normal behaviors.

The optimal feature selection executed through a combination of Support Vector Machine-Based Backward Feature Elimination (SVM-BFE), focusing on refining the feature space to retain the most informative and discriminative features, while eliminating the redundant and less contributive ones. This feature optimization is crucial for enhancing the model's predictive accuracy and computational efficiency, thereby contributing to a more responsive and effective intrusion detection system (IDS).

The architecture figure 3.1 employs a sophisticated multilabel classification model, utilizing ensemble learning with a Random Forest algorithm. This model is tailored to classify diverse intrusion types, such as DDoS, backdoor, ransomware, and normal behaviors, facilitating a comprehensive intrusion detection scope within IoT fog computing landscapes. The ensemble learning approach empowers the model with robustness and stability, aggregating insights from multiple decision trees to ensure reliable and accurate intrusion type predictions. In the MCID architecture, every component, from feature fusion and optimal selection to multilabel classification, is meticulously integrated and fine-tuned. This results in an advanced IDS that is capable of identifying and differentiating multiple intrusion types with enhanced precision, fostering a more secure and resilient environment within the complex and dynamic ecosystems of IoT fog computing.

**3.1. The Feature.** The basis for intricate and successful models in intrusion detection systems is provided by behavioral, temporal, and anomalous features, particularly in dynamic and complex environments such as IoT fog computing. These features enhance detection capabilities by offering various angles and dimensions for analyzing and comprehending the intricate patterns and subtleties of the data.

Activity patterns of network entities are known as behavioral features. They reveal the typical and atypical behavior of entities through describing interactions, action sequences, and state transitions. These characteristics enable models to pick up on intricate behavioral details that might point to malicious activity or security

lapses. These features include user behaviors, system interactions, and network communication patterns. Early detection of anomalies and intrusions is facilitated by an understanding of these features.

On the other hand, temporal features are necessary to record the time-related elements and sequences of network events. Timestamps, durations, activity frequency, and the intervals between events or actions are some examples of these features. In order to identify suspicious or unusual patterns and trends, intrusion detection models can use time-based features to analyze activity sequences and timings. A layer of security is added by identifying sophisticated and stealthy attacks that use time to go undetected through the analysis of temporal dynamics.

Anomaly Features draw attention to anomalies, outliers, and deviations from the norm in the data. Statistical measurements, frequency distributions, and outlier scores—which indicate data deviation—are some examples of these features. Detection models can detect unusual patterns and irregularities, even new ones, by integrating anomaly features. By identifying new threats, abnormal features aid in the adaptation and survival of intrusion detection systems.

**3.2. Optimal Feature Selection.** Feature selection is a crucial process in the realm of machine learning, especially in scenarios where the data dimensionality is large. It becomes even more significant when applied to sensitive domains like intrusion detection in IoT fog computing environments. This is because irrelevant or redundant features can lead to overfitting, poor performance, and reduced interpretability.

SVM is a supervised machine learning algorithm, primarily used for classification tasks. It works by finding the hyperplane that best separates the classes in the feature space. Given its capability to handle high-dimensional data and its effectiveness in finding margin-based distinctions, SVM is a suitable model for the complex and dynamic environment of IoT fog computing.

BFE is a recursive method that begins by training the model using all available features. Post training, it evaluates model performance and identifies the least significant feature. This feature is eliminated, and the model is retrained with the remaining features. This process is repeated until a desired number of features is reached or further elimination results in unacceptable performance degradation.

Support Vector Machine (SVM) [19] combined with Backward Feature Elimination (BFE) offers an effective methodology for optimal feature selection. Here's a comprehensive description of this method in the context of intrusion detection:

The SVM-BFE (Support Vector Machine-Backward Feature Elimination) process (see figure 3.2) starts with standardizing the data due to the sensitivity of SVM to the scale of input data and initializing the model with all available features. The initial SVM model, trained with all features, undergoes performance assessment using metrics such as accuracy, F1 score, or the area under the ROC curve, preferably using cross-validation methods on the training set. The least significant feature, which has the minimum impact on model performance, is then identified and eliminated from the feature set, and the model is retrained with the reduced feature set. This process of feature elimination and model retraining is iteratively carried out, with the model's performance being assessed at each step, until further removal degrades the model's performance significantly, or a predetermined number of features are left.

After the iterative process, the remaining features constitute the optimal feature set, which is expected to provide the best generalization on unseen data. The final model, built using this optimal feature set, undergoes validation against a separate test set to assess its generalization capability, ensuring its reliability and effectiveness in accurately classifying and predicting the target variables. This meticulous and iterative process aims to strike a balance between model simplicity and performance, enhancing the model's predictability and interpretability, crucial for applications such as intrusion detection in IoT fog computing environments.

**Algorithm.** Let  $X = x_1, x_2, \dots, x_m$  be the feature space where  $x_i$  represents each feature, and  $m$  is the total number of features.

Let  $Y$  be the set of class labels,  $Y = y_1, y_2, \dots, y_n$ .

Let  $D$  be the dataset represented as  $D(x^1, y^1), (x^2, y^2), \dots, (x^N, y^N)$ , where  $N$  is the number of samples.

*SVM Decision Function.* SVM finds the optimal hyperplane that separates the different classes in the feature space. The decision function of a linear SVM can be represented as:  $f(x) = w^T x + b$  where  $f(x)$  is the decision function.  $w$  is the weight vector.  $x$  is the input feature vector.  $b$  is the bias term.

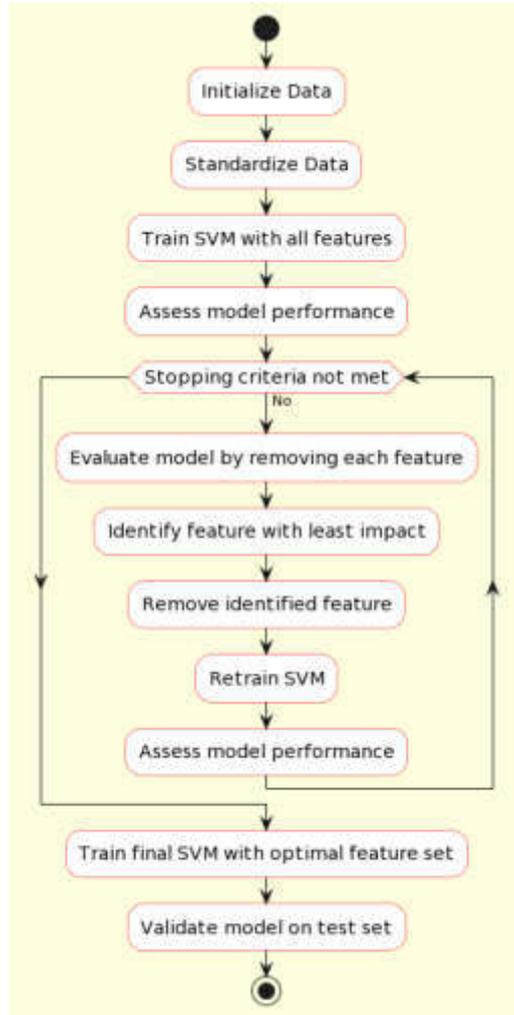


Fig. 3.2: SVM-BFE flow diagram

*Backward Feature Elimination (BFE)*. In the Backward Feature Elimination process, the least important feature is iteratively removed until an optimal subset of features is obtained. This can be represented as: Start with the full feature set:

$$F = x_1, x_2, \dots, x_m \quad (3.1a)$$

For each feature  $x_i$  in  $F$ , evaluate the SVM model, removing  $x_i$  from the feature set and note the performance. Identify the feature whose removal has the least impact on (or improves) the model's performance:

$$x_{remove} = \operatorname{argmin}_{(x_i \in F)} P(F - x_i) \quad (3.2a)$$

where  $P(F - x_i)$  represents the performance of the model trained with the feature set  $F$  excluding the feature  $x_i$ . Update the feature set by removing the identified feature: Eq 3.3

$$F = F - \{x_{remove}\} \quad (3.3a)$$

Repeat steps 2-4 until a stopping criterion is met, such as a pre-defined number of features or a performance threshold.

*Objective.* The objective is to find the optimal subset of features,  $F^*$ , which maximizes the performance of the SVM model:

$$F^* = \operatorname{argmax}_F P(F) \quad (3.4a)$$

*Stopping Criteria.* The iterative process stops when either:

The addition of any feature does not improve the model's performance. A predefined number of features is reached.

*Final Model.* The final optimal SVM model with the selected features is given by:

$$f^*(x) = w^{*T}x + b^* \quad (3.5a)$$

where  $w^*$  and  $b^*$  are the optimal weight vector and bias term obtained after training the SVM with the optimal feature subset  $F^*$ .

**3.3. Multiclass classification by Random Forest.** IoT Fog Computing forms a pivotal network architecture, allowing data processing, storage, and applications to operate closer to the end users along the cloud-to-thing continuum. Its dynamic and intricate nature makes it susceptible to a variety of cyber threats, including DDoS, ransomware, injections, backdoors, and other malicious activities, highlighting the critical need for robust intrusion detection systems (IDS) [20]. In this context, Random Forest emerges as a potent classification algorithm, proficient in predicting varying intrusion types and discerning normal behavior from malicious. The flow of the multiclass classification using RF [21] has been shown in figure 3.3.

Random Forest is an ensemble learning method, combining multiple decision trees [22] to construct a 'forest' that collaborates to render more accurate and stable predictions. This algorithm is inherently suited for multiclass classification tasks, making it apt for identifying diverse intrusion types like DDoS [23], ransomware [24], injection [25], backdoor [26], and normal activities within. IoT fog computing ecosystems [27]

Figure 3.3 depicts the workflow of multiclass intrusion detection module of the MCID that begins with loading a raw dataset and then extracts behavioral, temporal, anomaly features through feature engineering to form comprehensive feature matrix  $X$ ; next is preprocessing data devoid of missing values or normalization being processed as input required for Data is pre-processed, then split into training data and testing data. The former is used to train a Random Forest model which employs an ensemble approach by constructing numerous decision trees at different depths; the result of classifying the example chosen for given features are aggregated from all these trees so that output takes one with highest frequency as answer. Each tree in the model gives a vote to classify incoming classification data, which culminates in a majority decision. The final class label is then decided by potential classes such as Normal, DDoS (distributed denial-of-service), Ransomware, Injection or Backdoor. This procedure ends up with hyper-parameter tuning and optimizing, to raise as high the model accuracy possible so that the intrusion detection system can be made more reliable and effective. Detailed description of each step involved in this module follows:

**Feature Engineering and Preprocessing.** In the realm of intrusion detection, feature engineering is a foundational step, where the raw data is transformed into a structured format. It includes the extraction of relevant features like behavioral, temporal, and anomaly features, followed by preprocessing tasks such as normalization, encoding, and handling missing values.

Let  $D$  be the raw dataset, and let  $X = x_1, x_2, \dots, x_n$  represent the set of extracted features after feature engineering, where  $n$  is the number of features.

*Preprocessing.* Normalization:

$$x_{i,\text{norm}} = \frac{x_i - \min(x_i)}{\max(x_i) - \min(x_i)} \quad (3.6a)$$

where  $x_{i,\text{norm}}$  is the normalized value of feature  $x_i$ .

*Training the Model.* The preprocessed data is utilized to train the Random Forest model. The algorithm constructs multiple decision trees during training, each deliberating on a random subset of the features, and thus, learning varied aspects of the data. Let  $T = t_1, t_2, \dots, t_m$  be the set of decision trees in the Random Forest, where  $m$  is the number of trees. Each tree  $t$  is constructed using a subset of the feature set  $X' \subset X$  and a subset of the training dataset  $D' \subset D$ .

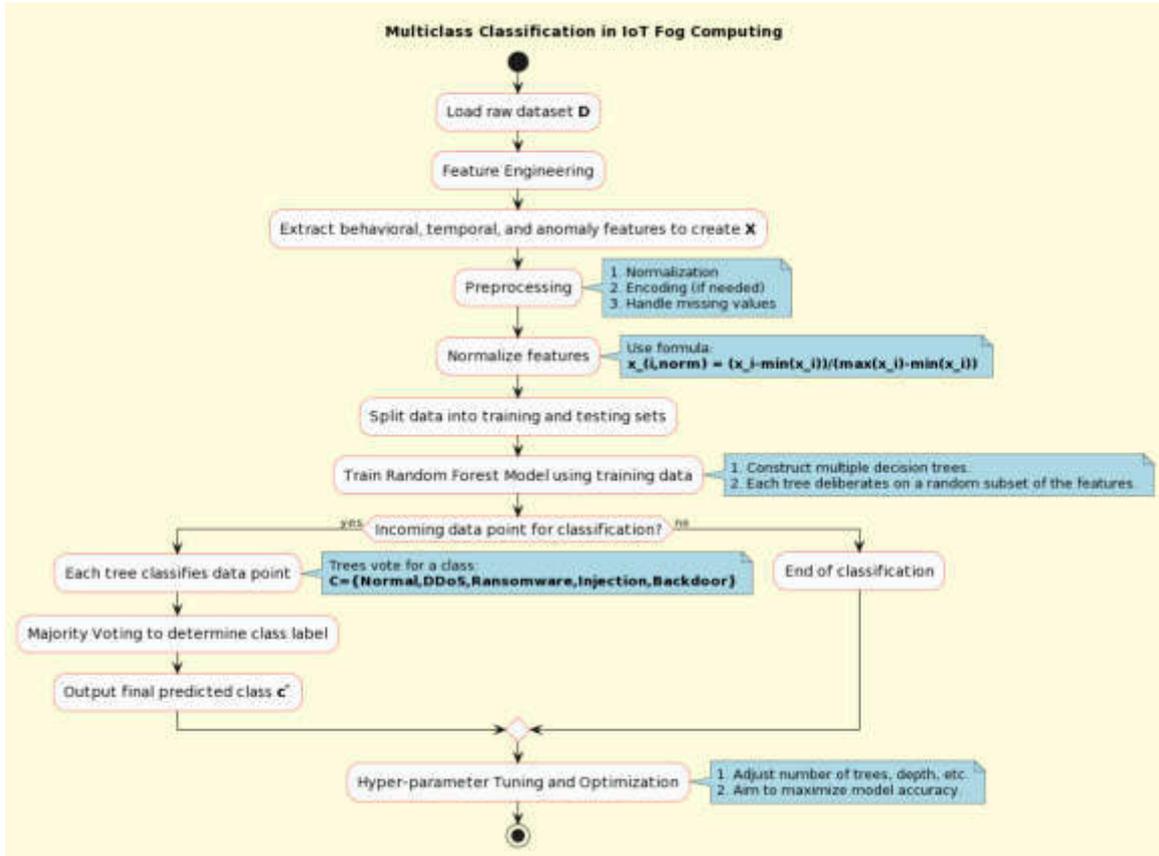


Fig. 3.3: Flow diagram of multiclass classification

*Classification and Prediction.* For every incoming data point, each tree in the ‘forest’ makes an individual decision, classifying the data point as either normal, DDoS, ransomware, injection, or backdoor. Subsequently, the Random Forest algorithm applies a majority voting mechanism to finalize the predicted class label, ensuring the collective wisdom of the ‘forest’ is reflected in every prediction [28]. Each tree  $t$  in votes  $T$  for a class  $c$  in the set of classes  $C = \text{Normal, DDoS, Ransomware, Injection, Backdoor}$ . The final predicted class  $\hat{c}$  is determined by majority voting:  $\hat{c} = \arg \max_{c \in C} \sum_{t \in T} 1(t(x) = c)$  where  $1$  is the indicator function, and  $t(x)$  is the class predicted by tree  $t$  for input  $x$ .

*Hyper-parameter Tuning and Optimization.* The model undergoes hyper-parameter tuning to refine its performance. Parameters like the number of trees, maximum depth of the trees, and minimum samples split are optimized to enhance the model’s predictive accuracy and generalization capability [29]. Let  $\Theta$  represent the set of hyperparameters of the Random Forest model, such as the number of trees and maximum depth. Hyperparameters tuning aims to find the optimal set of hyperparameters  $\Theta^*$  that maximize a given performance metric, say accuracy:

$$\Theta^* = \operatorname{argmax}_{\Theta} \operatorname{Accuracy}(\Theta) \quad (3.7a)$$

**4. Experimental Study.** The experimental study focuses on the evaluation of “Enhanced Feature Optimization for Multiclass Intrusion Detection in IoT Fog Computing Environments” (MCID) using data from the UNSW-NB15 [30] dataset. Balanced samples for each intrusion type and benign records were extracted, each consisting of 25,000 entries for BENIGN, DDOS, RANSOMWARE, and INJECTION. These were systematically divided into training and test subsets with 18,750 records designated for training and 6,250 for testing

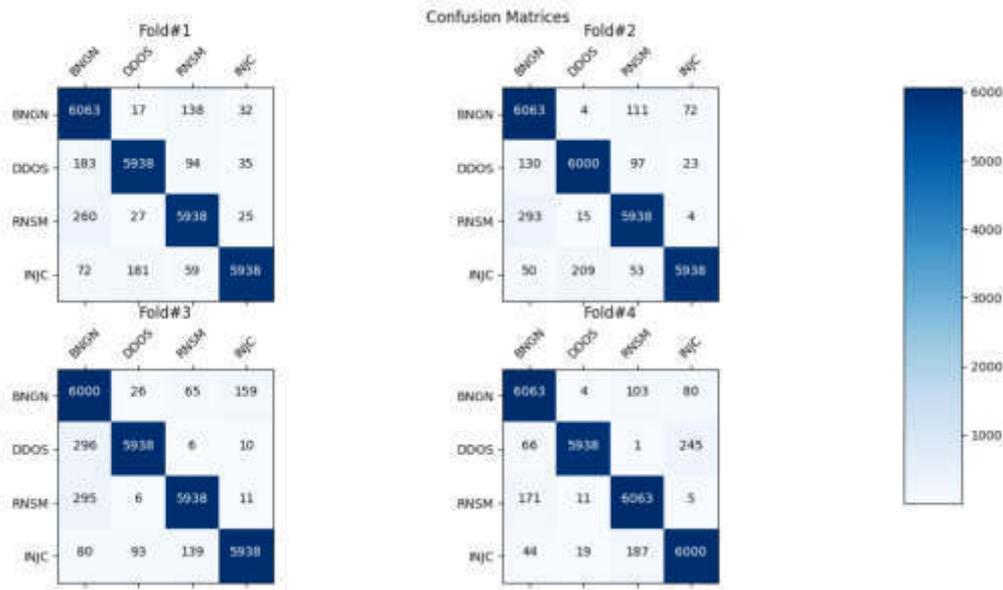


Fig. 4.1: Flow diagram of multiclass classification

across each label.

To validate the model’s performance, a four-fold cross-validation method was applied, ensuring consistent and reproducible results across various data partitions. The implementation leveraged Python and its associated libraries, highlighting the capabilities of this programming language in handling such sophisticated tasks.

The effectiveness and relevance of MCID were assessed in comparison to two contemporary models. These include "Lightweight Intrusion Detection Model of the Internet of Things with Hybrid Cloud-Fog Computing" (LIDM) [18] and "Combined Ensemble Intrusion Detection Model using Deep learning with Feature Selection for Fog Computing Environments" (CEIDM) [15]. This comparative analysis offers insights into the model’s capabilities, strengths, potential areas for enhancement, and its position in the broader spectrum of intrusion detection in IoT fog computing environments.

**4.1. Performance Analysis.** We examine the evaluation outcomes of a thorough 4-fold cross-validation exercise on the MCID, LIDM, and CEIDM datasets in the "Performance Analysis" section. To comprehend classification behavior and class predictions, we make use of confusion matrices. This performance metrics analysis teaches readers about recall, f-measure, specificity, precision, and more. The performance of the model is evaluated by looking at both aggregated macro metrics and individual class-wise metrics. The model’s performance is given by the overall accuracy metric. This part provides readers with a comprehensive overview of the model’s performance on various datasets, including its advantages and disadvantages. **MCID:** The provided confusion matrices in figure 4.1 represent the performance of a classification model across four folds of cross-validation on the MCID dataset. In each matrix, diagonal values depict correct predictions, while off-diagonal entries indicate misclassifications. Across the four folds, the model consistently demonstrates strong performance, particularly in the "DDOS" class, with minimal misclassifications. Notably, there are occasional misclassifications between "BENIGN" and the other classes, especially "RANSOMWARE" and "INJECTION". Despite these, the overall high true positive rates across classes suggest the model’s robustness on the MCID dataset.

In the provided metrics of the table 4.1, there is a consistent observation of high performance across all folds for the classes BNGN, DDOS, RNSM, and INJC. For Fold#1, the precision values range from 0.9217 for BNGN to 0.9847 for INJC. Similarly, for Fold#2, the precision values are slightly higher, ranging from 0.9276 for BNGN to 0.9836 for INJC. Fold#3 showcases the lowest precision for BNGN at 0.8994, but it also has the

Table 4.1: Performance Metric Values obtained from 4 fold cross validation performed on MCID

Fold	Metrics	BNGN	DDOS	RNSM	INJC
#1	Precision	0.9217	0.9635	0.9533	0.9847
#1	Recall	0.9701	0.9501	0.9501	0.9501
#1	F-measure	0.9453	0.9567	0.9517	0.9671
#1	Specificity	0.9725	0.988	0.9845	0.9951
#2	Precision	0.9276	0.9634	0.9579	0.9836
#2	Recall	0.9701	0.96	0.9501	0.9501
#2	F-measure	0.9484	0.9617	0.954	0.9666
#2	Specificity	0.9748	0.9878	0.9861	0.9947
#3	Precision	0.8994	0.9794	0.9658	0.9706
#3	Recall	0.96	0.9501	0.9501	0.9501
#3	F-measure	0.9287	0.9645	0.9579	0.9602
#3	Specificity	0.9642	0.9933	0.9888	0.9904
#4	Precision	0.9557	0.9943	0.9542	0.9479
#4	Recall	0.9701	0.9501	0.9701	0.96
#4	F-measure	0.9628	0.9717	0.9621	0.9539
#4	Specificity	0.985	0.9982	0.9845	0.9824

Table 4.2: Accuracy and Macro Metric Values obtained from 4 fold cross validation performed on MCID

Fold Id	Accuracy	Macro Precision	Macro Recall	Macro F-measure	Macro Specificity
#1	0.9775	0.9558	0.9551	0.9552	0.9850
#2	0.9788	0.9581	0.9576	0.9576	0.9859
#3	0.9763	0.9538	0.9526	0.9528	0.9842
#4	0.9813	0.9630	0.9626	0.9626	0.9875

highest precision for DDOS at 0.9794. Fold#4 stands out with a notably high precision for DDOS at 0.9943. Recall values remain fairly consistent across folds, mostly hovering around the 0.95 mark. The F-measure, which combines precision and recall, also shows commendable results with the highest being 0.9717 for DDOS in Fold#4. Specificity, which indicates the true negative rate, is highest for DDOS in Fold#3 and Fold#4, reaching up to 0.9982. Overall, the metrics indicate a strong performance with slight variances across folds and classes.

The table 4.2 provides an overview of the model's performance across four different folds. Accuracy values, which represent the proportion of correct predictions, consistently exhibit strong performance, ranging from 0.9763 in FOLD#3 to a peak of 0.9813 in FOLD#4. This high accuracy is further corroborated by the macro-level metrics. Macro Precision, which indicates the average precision across classes, demonstrates values above 0.95 for all folds, with FOLD#4 having the highest at 0.963. Similarly, Macro Recall, representing the average recall across classes, fluctuates around the 0.95 mark, with FOLD#4 slightly leading the pack. The Macro F-measure, a harmonic mean of precision and recall, echoes these findings with values very close to Macro Recall. Lastly, the Macro Specificity, which measures the true negative rate, consistently remains high across all folds, highlighting the model's capability to correctly identify negative cases, with the highest specificity seen in FOLD#4 at 0.9875. Overall, the model demonstrates robust and consistent performance across different folds.

**LIDM.** The confusion matrices presented in figure 4.2 provide insights into the performance of the model trained on the LIDM dataset across four folds of cross-validation. In each fold, the diagonal elements of the matrix represent the number of correct predictions for each class: BENIGN, DDOS, RANSOMWARE, and INJECTION. For instance, in Fold#1, the model correctly classified 5938 instances as BENIGN but misclassified 130 as RANSOMWARE and 175 as INJECTION. Similarly, 5813 DDOS instances were correctly

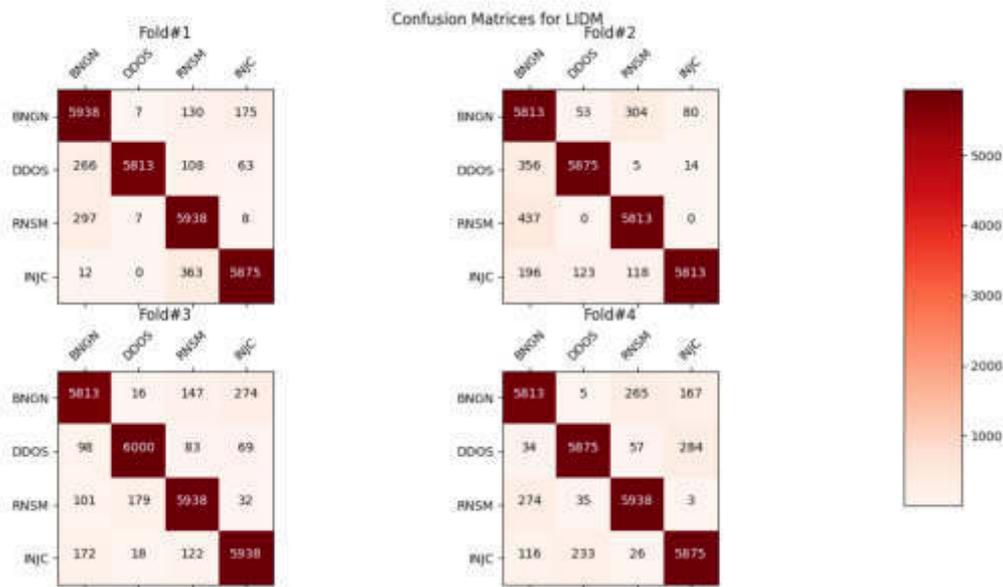


Fig. 4.2: Confusion matrices obtained from 4 fold cross validation performed on LIDM

Table 4.3: Performance Metrics by Fold and Category

	Metrics	BNGN	DDOS	RNSM	INJC
Fold#1	Precision	0.9117	0.9976	0.9081	0.9598
	Recall	0.9501	0.9301	0.9501	0.9400
	F-measure	0.9305	0.9627	0.9286	0.9498
	Specificity	0.9693	0.9993	0.9679	0.9869
Fold#2	Precision	0.8546	0.9709	0.9316	0.9841
	Recall	0.9301	0.9400	0.9301	0.9301
	F-measure	0.8907	0.9552	0.9308	0.9563
	Specificity	0.9473	0.9906	0.9772	0.9950
Fold#3	Precision	0.9400	0.9657	0.9440	0.9406
	Recall	0.9301	0.9600	0.9501	0.9501
	F-measure	0.9350	0.9629	0.9470	0.9453
	Specificity	0.9802	0.9886	0.9812	0.9800
Fold#4	Precision	0.9320	0.9556	0.9446	0.9283
	Recall	0.9301	0.9400	0.9501	0.9400
	F-measure	0.9310	0.9477	0.9474	0.9341
	Specificity	0.9774	0.9854	0.9814	0.9758

identified, but 108 of them were falsely predicted as RANSOMWARE. Each fold displays some variability in misclassifications across classes. For example, Fold#2 shows a noticeable increase in misclassifications of BENIGN as RANSOMWARE compared to Fold#1. On the other hand, in Fold#3 and Fold#4, the number of correctly predicted RANSOMWARE instances is consistent at 5938. Overall, the confusion matrices provide a comprehensive view of the model’s strengths and weaknesses across different data splits in the four-fold cross-validation.

In the table 4.3 four-fold cross-validation of the LIDM application, the performance metrics clearly exhibit its robustness in threat detection and classification. For instance, in Fold#1, DDOS detection exhibits an exceptional precision of 0.9976, indicating that 99.76% of the identified DDOS threats were accurate. This

Table 4.4: Accuracy and Macro Metric Values obtained from 4 fold cross validation performed on LIDM

Fold ID	Accuracy	Macro-Precision	Macro-Recall	Macro-F-measure	Macro-Specificity
#1	0.9713	0.944	0.9426	0.9429	0.9809
#2	0.9663	0.935	0.9326	0.9333	0.9775
#3	0.9738	0.948	0.9476	0.9476	0.9825
#4	0.9700	0.940	0.9400	0.9401	0.9800

remarkable precision is echoed in Fold#2 for the INJC category with a value of 0.9841. The recall values, which reflect the system's sensitivity, remain consistent, as evidenced by RNSM's recall of 0.9501 in both Fold#1 and Fold#3. Furthermore, the F-measure, which is a harmonized average of precision and recall, remains commendably high across all threats, such as the 0.9627 for DDOS in Fold#1 and 0.9563 for INJC in Fold#2. Specificity, indicating the system's ability to correctly identify non-threats, is notably high in Fold#1 for DDOS at 0.9993 and for INJC in Fold#2 at 0.995. All these metrics emphasize the reliability and efficacy of the LIDM application in the realm of cyber threat detection and classification.

In the table 4.4 four-fold cross-validation evaluation of the LIDM application, the metrics consistently demonstrate a high level of performance across all folds. The accuracy of the application remains stellar, ranging between 96.63% in Fold#2 and 97.38% in Fold#3. This trend of remarkable proficiency is also reflected in the Macro-Precision, with values spanning from 0.9353 in Fold#2 to 0.9476 in Fold#3. Similarly, the Macro-Recall, indicating the application's sensitivity in correctly identifying threats, consistently hovers around the 94% mark across all folds. The Macro-F-measure, a harmonized average of precision and recall, fortifies this trend by consistently staying above 93%. Lastly, the Macro-Specificity, a measure of the application's aptitude in correctly recognizing non-threats, remains impressive, reaching as high as 98.25% in Fold#3. All these metrics collectively underscore the effectiveness and reliability of the LIDM application in its operational domain.

**CEIDM.** The confusion matrices presented in figure 4.3 for CEIDM across four folds depict the classification performance for the categories: "BENIGN", "DDOS", "RANSOMWARE", and "INJECTION". Across the folds, the diagonals of the matrices, which represent the true positive classifications, show consistently high values. This indicates a generally high accuracy in classification for each category. However, there are evident misclassifications. For example, in Fold#1, "RANSOMWARE" has been occasionally mistaken as "DDOS", while "INJECTION" has a considerable number of false predictions across other categories. Similarly, in other folds, while the diagonal values remain dominantly high, indicating successful classifications, the off-diagonal values point out areas where the model struggled, resulting in misclassifications. These matrices offer valuable insights for further refinement of the classification model to enhance its accuracy across all categories.

The table 4.5 presents performance metrics of CEIDM, evaluated across four folds. These metrics include Precision, Recall, F-measure, and Specificity for four classes: BNGN, DDOS, RNSM, and INJC. For Fold#1, the model exhibits high precision for the BNGN class at 0.9747, while the DDOS class has the lowest precision of 0.8971. However, the recall for DDOS is higher than BNGN in this fold. This trend of varying precision and recall figures is observed across all folds. The F-measure, which combines precision and recall, also demonstrates varied performance among classes, with values mostly hovering above 0.9. Specificity, indicating the true negative rate, remains notably high across all classes and folds, often surpassing 0.97. By Fold#4, the model's precision for the RNSM class peaks at 0.9547. In order to sum up, the CEIDM model exhibits good performance for all metrics and classes, though there is some variation between the folds.

The table 4.6 presents an overview of the aggregate performance metrics for the CEIDM model over four folds. The metrics encapsulated include Accuracy, Macro-Precision, Macro-Recall, Macro-F-measure, and Macro-Specificity. Throughout all folds, the model consistently maintains an accuracy around 0.96, indicating a strong overall classification performance. The Macro-Precision and Macro-Recall, which provide an average performance measure across the classes, mostly hover in the low 0.93s to high 0.92s range, denoting that the model is both precise and sensitive in its predictions. The Macro-F-measure, which combines the precision and recall, mirrors these figures closely. Notably, the Macro-Specificity remains robust across all folds, predominantly exceeding 0.97, suggesting that the model is adept at correctly identifying negatives. In essence, CEIDM

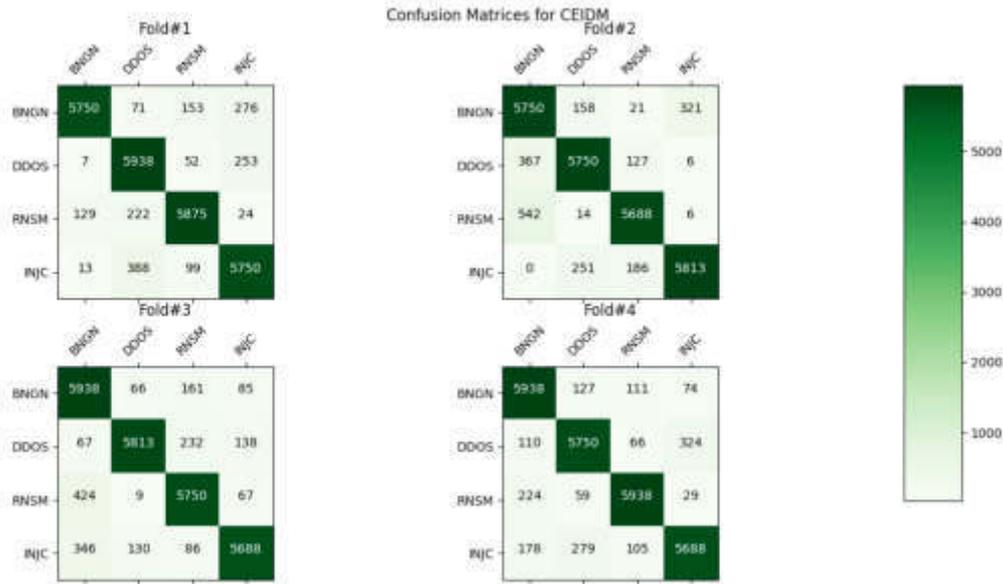


Fig. 4.3: Confusion matrices obtained from 4 fold cross validation performed on CEIDM.

Table 4.5: Performance Metrics by Fold and Metric Performance Metric Values obtained from 4 fold cross validation performed on CEIDM

Fold's	Metrics	BNGN	DDOS	RNSM	INJC
Fold#1	Precision	0.9747	0.8971	0.9508	0.9123
	Recall	0.9200	0.9501	0.9400	0.9200
	F-measure	0.9466	0.9228	0.9454	0.9161
	Specificity	0.9921	0.9637	0.9838	0.9705
Fold#2	Precision	0.8635	0.9315	0.9445	0.9458
	Recall	0.9200	0.9200	0.9101	0.9301
	F-measure	0.8909	0.9257	0.9270	0.9379
	Specificity	0.9515	0.9774	0.9822	0.9822
Fold#3	Precision	0.8765	0.9659	0.9231	0.9515
	Recall	0.9501	0.9301	0.9200	0.9101
	F-measure	0.9118	0.9477	0.9215	0.9303
	Specificity	0.9554	0.9891	0.9745	0.9845
Fold#4	Precision	0.9206	0.9252	0.9547	0.9302
	Recall	0.9501	0.9200	0.9501	0.9101
	F-measure	0.9351	0.9226	0.9524	0.9200
	Specificity	0.9727	0.9752	0.9850	0.9772

Table 4.6: Accuracy and Macro Metric Values obtained from 4 fold cross validation performed on CEIDM

Fold ID	Accuracy	Macro-Precision	Macro-Recall	Macro-F-measure	Macro-Specificity
#1	0.9663	0.934	0.9325	0.9327	0.9775
#2	0.9600	0.921	0.9200	0.9204	0.9733
#3	0.9638	0.929	0.9276	0.9278	0.9759
#4	0.9663	0.933	0.9326	0.9325	0.9775

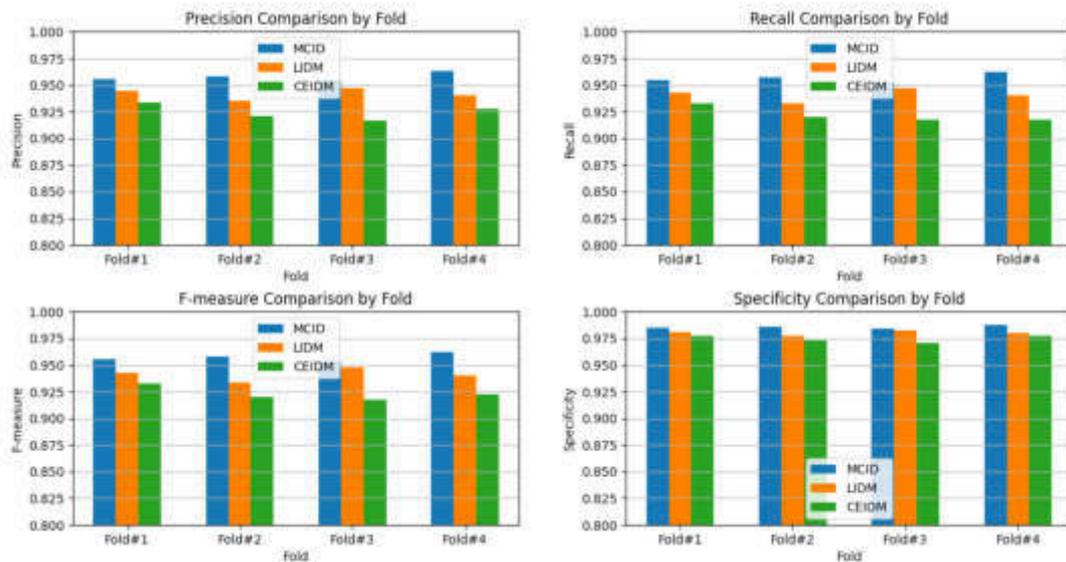


Fig. 4.4: Comparative analysis of precision, recall, f-measure, and specificity

consistently delivers reliable and robust performance metrics across the four evaluation folds.

**4.2. Comparative Analysis.** This section provides evaluation comparing the proposed MCID model with its contemporaries, LIDM and CEIDM. Through a meticulous examination of performance across vital metrics such as precision, recall, F-measure, specificity, and accuracy, along with their macro counterparts, insights are drawn regarding the inherent strengths, potential pitfalls, and distinguishing characteristics of each model. This analysis seeks to offer a comprehensive perspective on the standing of MCID within the contemporary landscape.

According to comparative analysis presented in figure 4.4, MCID consistently demonstrates commendable performance across all metrics and folds, indicating its robustness and generalizability. Notably, it exhibits high specificity, especially in the initial folds, ensuring a low false positive rate which is crucial in many applications. Additionally, its balanced results across precision, recall, and F-measure highlight its capability to provide reliable classification without significant trade-offs. Even in scenarios where it doesn't lead, MCID remains a close competitor, showcasing its potential as an effective and versatile method in comparison to LIDM and CEIDM.

The bar graphs in figure 4.5 provide insight into the performance of three methods (MCID, LIDM, and CEIDM) across four different folds. At a glance, MCID consistently offers the highest accuracy across all folds, with FOLD#4 being its best at 0.9813. While LIDM and CEIDM often hover close in metrics, LIDM slightly outperforms CEIDM in accuracy for most folds. However, in terms of Macro Precision, Recall, and F-measure, MCID again generally surpasses the other two, solidifying its position as the most effective model among the three, especially in FOLD#4. For Macro Specificity, all methods perform quite commendably, with minimal variance across folds. Yet, MCID again edges out with the highest value in FOLD#4. Overall, the MCID method proves superior, with FOLD#4 consistently representing the highest metric values for this method.

A comparative analysis table 4.7 summarizes the average performance improvement of MCID over LIDM and CEIDM with respect to four key metrics. The list presents percentages (showing improvements) along with their standard deviations. Compared with LIDM and CEIDM, the model is very reliable in terms of detailing every step along the way. Every time it was tested for its ability to forecast results, on average a 0.8425 % improvement occurred when compared with LIDM and an even larger increase of 1.4975 % over that achieved by CEIDM. Furthermore, in Macro Precision it is better than LIDM by 1.685 % and CEIDM by

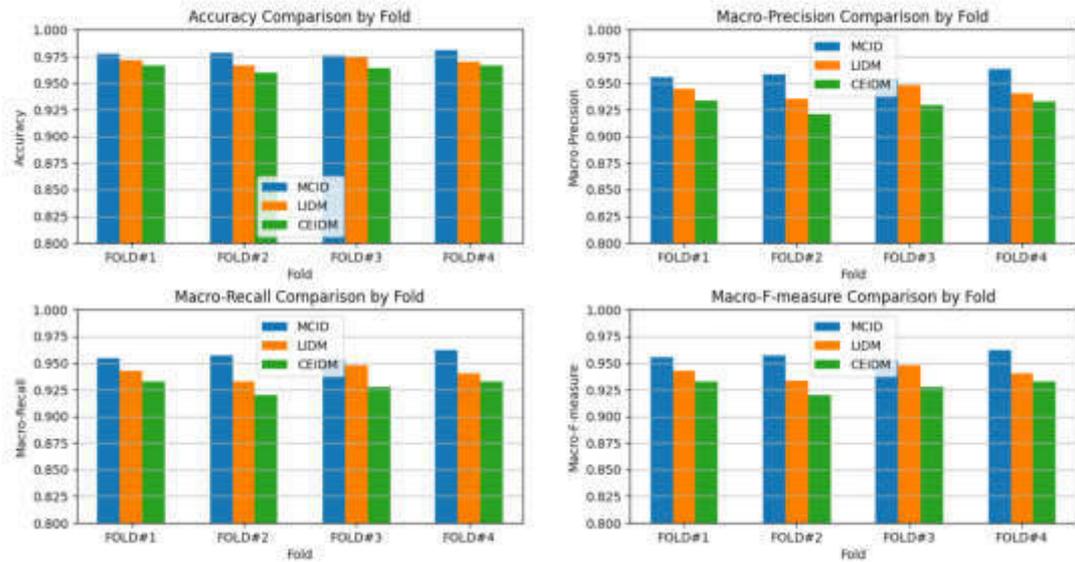


Fig. 4.5: Comparative analysis of Accuracy and macro values of the precision, recall, and f-measure

Table 4.7: comparative table of improvement rate

Metric	Improvement vs LIDM (%)	Improvement vs CEIDM (%)
Accuracy	0.8425	1.4975
Macro Precision	1.685	3.08
Macro Recall	1.735	3.1125
Macro F-measure	1.71	3.1
Macro Specificity	0.55	0.975

3.08 %, while in terms of Macro Recall the latter figure comes to as much as 1.735 % over against those for respectively approximately definitive idiomatic expressions (LIDM) and common sense fine-grained Macro F-measure’s general performance Is MCID 1.71% ahead and 3.1 % higher over the two other models respectively, which suggests that this is a balanced improvement in both precision and sensitivity of classifications at its peak value for each parameter (F max). For Macro Specificity, MCID retains a small edge over LIDM (240.86P compared to 239.57 P) and an even greater margin over CEIDM at 1 point of distortion error or more—a whopping improvement in both cases! These figures together attest to MCID’s strength and its prowess at helping unleash the full potential of intrusion detection in IoT fog computing environments.

**5. Conclusion.** In sum, the MCID framework represents an important step forward in IoT fog computing security. It fills the critical gaps that traditional security measures are often unprepared to fill, dealing with fog computing environments which by their nature tend to be complex and constantly evolving. With a rich set of features and sophisticated machine learning techniques, MCID is offered as an efficient solution to the problems of intrusion detection. By its detailed architecture, the system can fully exploit behavioral, temporal and anomaly characteristics, optimizing them through SVM-BFE. This leads to a streamlined feature set that improves the system’s accuracy and speed. Also, using a Random Forest algorithm for multilabel classification helps MCID precisely identify various kinds of intrusion targets. Cross-validation results further show that the MCID framework has high efficacy across many different metrics, making it a potent weapon in comprehensive security. Based on the precision, recall rate. F-measure and specificity values it achieves accuracy in threat detection and classification while minimizing false positives to grant IoT devices a high level of reliable protection

under fog computing networks. In short, the design of MCID is an important step in making fog computing secure. Its novel methodology and early successes open the gateway for more research and development on this front, foreshadowing a model that is scalable as well as flexible enough to conform with changes in the threat environment created by IoT.

## REFERENCES

- [1] M. LOMBARDI, F. PASCALE, AND D. SANTANIELLO, *Internet of things: A general overview between architectures, protocols and applications*, Information, 12(2), 2021, pp. 87.
- [2] S. QABIL, U. WAHEED, S. M. AWAN, Y. MANSOOR, AND M. A. KHAN, *A survey on emerging integration of cloud computing and internet of things*, In 2019 International Conference on Information Science and Communication Technology (ICISCT), IEEE, 2019, pp. 1-7.
- [3] A. OMETOV, O. L. MOLUA, M. KOMAROV, AND J. NURMI, *A survey of security in cloud, edge, and fog computing*, Sensors, 22(3), 2022, pp. 927.
- [4] B. SUDQI KHATER, A. W. B. A. WAHAB, M. Y. I. B. IDRIS, M. A. HUSSAIN, AND A. A. IBRAHIM, *A lightweight perceptron-based intrusion detection system for fog computing*, Applied Sciences, 9(1), 2019, pp. 178.
- [5] K. SADAF AND J. SULTANA, *Intrusion detection based on autoencoder and isolation forest in fog computing*, IEEE Access, 8, 2020, pp. 167059-167068.
- [6] P. KUMAR, G. P. GUPTA, AND R. TRIPATHI, *A distributed ensemble design based intrusion detection system using fog computing to protect the internet of things networks*, Journal of Ambient Intelligence and Humanized Computing, 12, 2021, pp. 9555-9572.
- [7] K. KALAIVANI AND M. CHINNADURAI, *A Hybrid Deep Learning Intrusion Detection Model for Fog Computing Environment*, Intelligent Automation & Soft Computing, 30(1), 2021.
- [8] F. A. ZWAYED, M. ANBAR, Y. SANJALAWA, AND S. MANICKAM, *Intrusion Detection Systems in Fog Computing—A Review*, In Advances in Cyber Security: Third International Conference, ACeS 2021, Penang, Malaysia, August 24–25, 2021, Revised Selected Papers 3, Springer Singapore, 2021, pp. 481-504.
- [9] H. A. AFOLABI AND A. ABURAS, *Proposed back propagation deep neural network for intrusion detection in internet of things fog computing*, Int J, 9(4), 2021, pp. 464-469.
- [10] M. RAMKUMAR, *Support vector machine based intrusion detection system in fog computing*, ICTACT Journal on Data Science and Machine Learning, March 2021, Volume: 02, Issue: 02, pp. 16-164.
- [11] V. CHANG, L. GOLIGHTLY, P. MODESTI, Q. A. XU, L. M. T. DOAN, K. HALL, S. BODDU, AND A. KOBUSIŃSKA, *A survey on intrusion detection systems for fog and cloud computing*, Future Internet, 14(3), 2022, pp. 89.
- [12] A. WU, S. TU, M. WAGAS, Y. YANG, Y. ZHANG, AND X. BAI, *Intrusion Detection System Using a Distributed Ensemble Design Based Convolutional Neural Network in Fog Computing*, Journal of Information Hiding and Privacy Protection, 4(1), 2022, pp. 25.
- [13] O. A. ALZUBI, J. A. ALZUBI, M. ALAZAB, A. ALRABEA, A. AWAJAN, AND I. QIQIEH, *Optimized machine learning-based intrusion detection system for fog and edge computing environment*, Electronics, 11(19), 2022, pp. 3007.
- [14] J. SAJID, K. HAYAWI, A. W. MALIK, Z. ANWAR, AND Z. TRABELSI, *A Fog Computing Framework for Intrusion Detection of Energy-Based Attacks on UAV-Assisted Smart Farming*, Applied Sciences, 13(6), 2023, pp. 3857.
- [15] K. KALIYAPERUMAL, C. MURUGAIYAN, D. PERUMAL, G. JAYARAMAN, AND K. SAMIKANNU, *Combined Ensemble Intrusion Detection Model using Deep learning with Feature Selection for Fog Computing Environments*, Acta Scientiarum. Technology, 45, 2023.
- [16] D. MOHAMED, AND O. ISMAEL, *Enhancement of an IoT hybrid intrusion detection system based on fog-to-cloud computing*, Journal of Cloud Computing, 12(1), 2023, pp. 1-13.
- [17] C. A. DE SOUZA, C. B. WESTPHALL, AND R. B. MACHADO, *Intrusion detection with Machine Learning in Internet of Things and Fog Computing: problems, solutions and research*, Sociedade Brasileira de Computação, 2023.
- [18] G. ZHAO, Y. WANG, AND J. WANG, *Lightweight Intrusion Detection Model of the Internet of Things with Hybrid Cloud-Fog Computing*, Security and Communication Networks, 2023, 2023.
- [19] D. A. PISNER, AND D. M. SCHNYER, *Support vector machine*, In Machine learning, Academic Press, 2020, pp. 101-121.
- [20] A. KHRAISAT, I. GONDAL, P. VAMPLEW, AND J. KAMRUZZAMAN, *Survey of intrusion detection systems: techniques, datasets and challenges*, Cybersecurity, 2(1), 2019, pp. 1-22.
- [21] S. J. RIGATTI, *Random forest*, Journal of Insurance Medicine, 47(1), 2017, pp. 31-39.
- [22] L. ROKACH AND O. MAIMON, *Decision trees*, Data mining and knowledge discovery handbook, 2005, pp. 165-192.
- [23] R. VISHWAKARMA AND A. K. JAIN, *A survey of DDoS attacking techniques and defence mechanisms in the IoT network*, Telecommunication systems, 73(1), 2020, pp. 3-25.
- [24] P. O'KANE, S. SEZER, AND D. CARLIN, *Evolution of ransomware*, IET Networks, 7(5), 2018, pp. 321-327.
- [25] R. D. JAGER, L. P. AIELLO, S. C. PATEL, AND E. T. CUNNINGHAM JR., *Risks of intravitreal injection: a comprehensive review*, Retina, 24(5), 2004, pp. 676-698.
- [26] Y. LI, Y. JIANG, Z. LI, AND S.-T. XIA, *Backdoor learning: A survey*, IEEE Transactions on Neural Networks and Learning Systems, 2022.
- [27] H. F. ATLAM, R. J. WALTERS, AND G. B. WILLS, *Fog computing and the internet of things: A review*, Big Data and Cognitive Computing, 2(2), 2018, pp. 10.
- [28] Y.-Y. SONG AND L. U. YING, *Decision tree methods: applications for classification and prediction*, Shanghai Archives of

- Psychiatry, 27(2), 2015, pp. 130.
- [29] T. YU AND H. ZHU, *Hyper-parameter optimization: A review of algorithms and applications*, arXiv preprint arXiv:2003.05689, 2020.
- [30] N. MOUSTAFA AND J. SLAY, *UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set)*, In 2015 Military Communications and Information Systems Conference (MilCIS), Nov 2015, pp. 1–6.

*Edited by:* Anil Kumar Budati

*Special issue on:* Soft Computing and Artificial Intelligence for wire/wireless Human-Machine Interface

*Received:* Oct 4, 2023

*Accepted:* Jan 13, 2024



## QUALITY ENHANCEMENT WITH FRAME-WISE DLCNN USING HIGH EFFICIENCY VIDEO CODING IN 5G NETWORKS

VIJAYA SARADHI DOMMETI\*, M. DHARANI†, K. SHASIDHAR‡, Y DASARATHA RAMI REDDY§ AND T. VENKATAKRISHNAMOORTHY¶

**Abstract.** In the present situation, applications related to multimedia are discovered to be comfortable with the use of video. The number of end consumers who use video continues to rise every day. People are presently searching for videos with better quality among the ones that are currently available there. This results in the launch and dissemination of HD (high definition) videos. Ultra high-definition (UHD) videos are becoming more and more popular as a result of this advancement and need. However, as video communication keeps expanding, there is an upsurge in network traffic because of the limited bandwidth, especially among smart cities. Different advancement codecs have been suggested to deal with the data stream to overcome this hazardous circumstance. However, the fact that modern UHD videos have huge amounts of data makes the available codecs even more complicated. UHD videos can be processed with the latest improvement codec, H.265/High-efficiency video coding (HEVC). Nevertheless, it is impacted by increased power consumption and intricate calculations. Limitations in the codec's functionality confine its use to specific applications, preventing its application in wireless, mobile, or portable settings. Hence, this research concentrates on implementing frame-level quality enhancement through a deep learning network known as FQE-Net. The deep learning convolutional neural network (DLCNN) is specifically crafted to manage films with resolutions up to 16K. Its primary objectives include reducing complexity, minimizing artifacts, enhancing the efficiency of the HEVC codec, and compacting energy consumption. To achieve superior efficiency, it is imperative to replace the DWT transforms within the HEVC codec with a DLCNN model. Additionally, incorporating the Content Block Search Algorithm for Motion Estimation and Compensation, alongside filtering techniques like Sample Adaptive Filter and Deblocking Filter, becomes essential. The simulation results showed that the suggested FQE-Net performed better than the conventional techniques.

**Key words:** Convolutional neural networks for deep learning, high-efficiency video codec, and ultra high definition.

**1. Introduction.** End customers expect films with good resolution these days. Consequently, the demands of the end user are being met with difficulty by the multimedia firm and researchers [1]. The average user doesn't care about the technical details of the movie sequence; all that matters to them is watching a video of the highest quality without any buffering [2]. Because of this, a poll in [3] indicates that it is challenging to use present technology to offer a video of competitively good quality at a lower bit-rate and a higher enhancement ratio. Nevertheless, the surge in demand for superior-quality products is propelled by several factors, including the introduction of expanded communication bandwidths in networks such as LTE and 5G, the proliferation of smart devices, advancements in high-resolution display systems, market growth, elevated user standards, and innovations in technology. These days, it's common to access and upgrade 4K/8K-UHD video contents for distant applications [5]. However, as image quality rises to UHD, the previous H.264/AVC codec's coding efficiency runs out of bandwidth. These days, it's common to access and upgrade 4K/8K-UHD video contents for distant applications [5]. However, as image quality rises to UHD, the previous H.264/AVC codec's coding efficiency runs out of bandwidth. A new architecture is proposed to address the erotic scenario and make HEVC appropriate for both future UHD videos and already available HD videos [6]. The suggested codec is made with minimal resource consumption, minimal computational complexity, optimal latency time, and a lower bit-rate combined with better video quality. Video is becoming increasingly important with the introduction of new electronic devices like smartphones, HDTVs, multimedia systems, video surveillance, and more [7, 8]. These devices carry out several functions, including high-definition video conferencing, web browsing, sharing

\*Department of ECE, Malineni Lakshmaiah womens Engineering college, Guntur, A.P., India. ([saradhi.gitam@gmail.com](mailto:saradhi.gitam@gmail.com))

†Department of ECE, Mohan Babu University, Tiruapthi, Andhra Pradesh, India ([ddharani405@gmail.com](mailto:ddharani405@gmail.com))

‡Department of ECE, Lords Institute of Engineering & Technology, Hyderabad, India ([shashignitc2015@gmail.com](mailto:shashignitc2015@gmail.com))

§Department of CSE, Chaitanya Bharathi Institute of Technology, Proddatur, A.P., India ([dasradh@gmail.com](mailto:dasradh@gmail.com))

¶Department of ECE, Sasi Institute of Technology & Engineering, Tadepalligudem, AP, India ([tvkmsvu@gmail.com](mailto:tvkmsvu@gmail.com))

social entertainment videos, and web monitoring transmission and distribution of video materials. One of the answers to the Digital Era [10], which has greatly impacted most industries, including communications, arts, entertainment, marketing, and media, is disruptive technology. This risky evolution is a byproduct of the modern period, which is distinguished by a dramatic shift from analog to digital technology. The speed at which digital content—including music, video, and data—is produced and accessed dictates how fast digital gadgets can process data. It is entirely up to us to decide what, if any, comes next. High-quality digital video has special opportunities and problems for analysis and visualization due to its extensive applications in domains such as storing data, internet streaming, monitoring, broadcasting, and conference calling [11].

1. FQE-Net's implementation to improve the video quality of HEVC standards based on H.265.
2. Using the DLCNN model will reduce complexity, energy compaction, and artifacts, and offer enhanced HEVC codec efficiency for handling videos up to 16K resolution.
3. The Motion Estimation and Compensation algorithms use the Content Block Search Algorithm; filtering techniques like the Sample Adoptive Filter and Deblocking Filter are also used.

The authors of [12] talked about using the FPS (Fast Predictive Search) algorithm for diamond searches. In this case, adding an initial search prediction is necessary. The search point in the first search step of our suggested technique is initially upgraded [13]. Here, the algorithm can locate the fundamental search point precisely and avoid the ineffective global search process. To reduce complexity, PIME (Predictive Integer Motion Estimation) is used in conjunction with a joint architecture algorithm to build an efficient high-motion estimation design. Based on analysis of potential search directions, this approach notably reduces the number of search points. The interpolation filtering procedure uses a processing unit and FME (Fractional Motion Estimation) [15]. For the computations, scheduling using cascade form is used for fractional motion estimation and integer motion estimation. By decreasing BW, gate count, and memory size, Z-scan [16] using indexed addressing simplifies the cache controller. A multi-rate encoding approach was presented by the authors in [17]; it is utilized to lower the encoding difficulty with several spatially distinctive resolutions. Finding similarities between the block structures of the various resolutions recorded using the high-resolution approach is the first step in the process. The resulting block structure is used to accelerate the encoding process [18]. The encoding of a low-resolution representation is accelerated by using a prefabricated block structure. To achieve individualized HEVC encoding with RD performance (Rate-Distortion), greater attention to the video content is needed. Based on simulation results, it is possible to cut the encoding time of flat-resolution video by 50 % on average without compromising rate-distortion performance [19].

The authors of [20] presented the Intra Mode and Inter Prediction decision approach. You should apply both tactics. A prediction with depth LCU technique with any adaptation is used in the third method, which is a generalized version of the first two Fast Intra Mode Decision methods [21]. We have developed four efficient early termination options to help you wrap up the RDO procedure quickly. Encoding times can be shortened by 54.4 %–68.54 %, based on the performance penalty. Based on a quality analysis of the material, the Fast Intra Prediction for Screen material Coding technique was first presented in the study [22]-[23]. The size of the current CU is essentially determined by its per-pixel data rate. Every CU is a part of this class. We are expanding the range of data stored by nearby CUs and collocated CUs in order to increase speed even more [24]. According to the trial results, the recommended approach can encode video information with an average cost reduction of 44.92 percent and with the least amount of picture quality loss. The creators of [25]-[26] improved the HEVC's HM 16 tool was the reason for the process's good encoding efficiency and 1 % increase in the BD rate. Reference 27 outlines a method that streamlines the calculation and expedites reinforcement learning. The authors of [28] suggest using ABCO, or adaptive block coding, as a method of built-in prediction. Following enhancement, a volumetric bit rate The spatial relationship between close and distant blocks ensures a volumetric bit rate when it has been improved. For cost savings, using a method that adaptively nominates block and sub-block coding orders is advised [29]. In opposition to HEVC interfering with set blocking orders. The potential for improving predictions in the edge region using adaptive means is the Advocate method in this work. Utilizing the suggested algorithm, the HM9.2 software produced 1.3 % outcomes and 4.4 % bit savings. The authors of [30]-[31] used a motion estimation approach to compensate for frame rate, which lessens computational complexity. The recommended method produces clearer images than baseline techniques, increasing PSNR by 1.3 dB. One potential method for finding exact motion vectors with minimal processing complexity was the

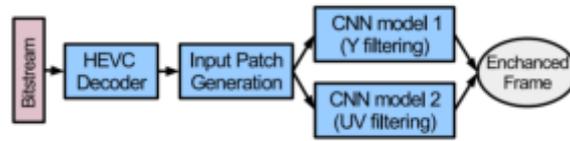


Fig. 2.1: FQE-Net proposal

Multi-Directional Motion assessment [32]. The Adaptive Motion Vector Smoothing Algorithm was proposed in the MVS phase as a different name for the AGMVS algorithm [33] in a way that modifies the motion vectors' failure by particular values in the in the unified framework. In light of the subjective judgment, the image is quite clear, but the objective examination resulted in a high PSNR.

The authors of [34] proposed using the Fast Coding Unit (FCU) choice for the HEVC encoding process, involving an offline approach for the Trained Supporting Vector Machines. By using the suggested approach, the proposed algorithm's complexity can be lowered by up to 48 % at a Bjontegard Delta Bit Rate of 0.48 [35]. 44 percent reduction in loss of RA (Random Access) setup using 0.62 % BD-BR. With a 41 percent decrease in LDB (Low Delay B) configuration and a 0.6 % BD-BR, the HEVC encoding complexity was efficiently reduced by a fast coding design approach, and the rate distortion optimization yielded competitive results. In reference [36], a new depth level and an inter-mode prediction method were introduced in an alternative to traditional SHVC. We first use interlayer correlation to predict potential depth ranges. The next step is to distinguish between squared and non-squared predictions within the depth candidate prediction mode [37, 38]. Gaussian distribution function is used for the early termination residual distribution. enhanced speed and efficiency of coding. The most computationally demanding activity and one that is extremely resilient to accuracy losses is the video encoding process with Motion estimation (ME) stage, which includes IME and FME techniques, as described by the authors in [39]. This paper provides an approximation of energy efficiency using motion estimation architectures that are supported by FME and IME [40].

**2. Proposed Method.** The proposed quality improvement system, depicted in Figure 2.1, employs the DLCNN-based filtering technique after fully decoding the current frame. Figure 1 illustrates this HEVC architecture modification. The two types of patches that are generated as the first step in the suggested technique filter the Y channel and ones that filter the U and V channels. The process then uses the two models to produce the increased chroma and luminance components,  $U_e$ ,  $V_e$  and  $Y_e$  respectively, for the enhanced frame,  $Y_e, U_e, V_e$ .

Block-based video coding is a widely used technique in the improvement process. Videos in this process can be categorized based on how many frames they contain. The most popular video encoding utilized by Block-based video coding is seen in Figure 2.2. Each frame in a particular frame sequence is divided into smaller parts using video coding blocks, from which the elements of frame 1 can be automatically predicted. Fundamentally, the motion estimation block generates a prediction before the application of the motion compensation block, or alternatively, the motion compensation block initiates an inter-prediction prior to the motion estimation block. Conversely, inter-prediction empowers the initial block of the slice algorithm to determine whether each frame should be intra-coded or not.

Prior to reaching the motion estimation and compensation block, the Motion Vectors undergo another round of entropy encoding. The present frames undergo subtraction from the Motion-Compensated frames (yielding the residual information) to generate the residual frame blocks. The rest of the frame blocks are quantized and changed before going into the entropy-based encoding stage. The identical set of projected data is available to both the encoder & the decoder. An encoder frequently incorporates a decoding loop as well, which uses the information at hand to piece together the original frames. In Images 2 and 3, dashed lines indicate the decoding blocks. Second, before the quantized data can be decoded, the decoder must pass it through inverse transformation blocks and inverse quantization.

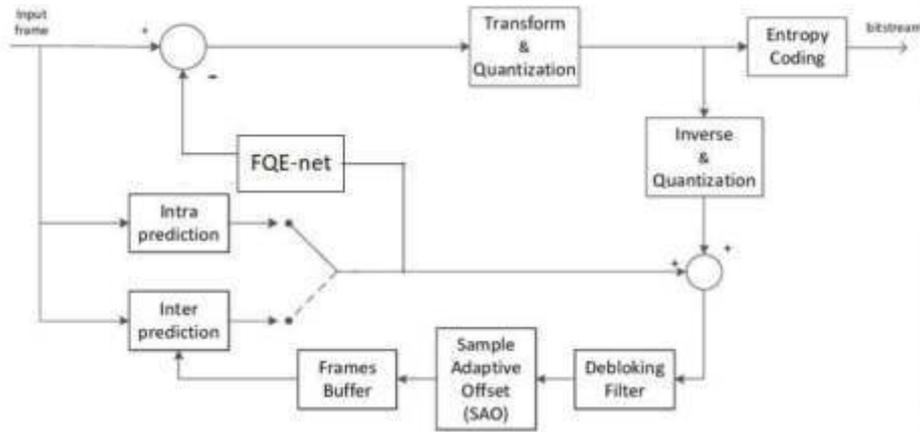


Fig. 2.2: The HEVC block diagram

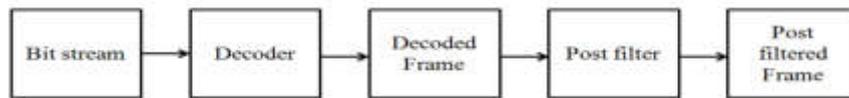


Fig. 2.3: Blocking artifacts reduction

**2.1. Looping filters in HEVC.** To decrease the effect of artifacts that obstruct highways, there are two general approaches that can be applied. They go by the labels post-filtering and loop-filtering, respectively. Figure 2.3 illustrates how the Deblocking filter is employed in post-filtering, which follows the decoder and makes use of the decoded parameters. It uses the display buffer to carry out its operations along with the main code loop. The frame needs to be filtered and encoded in the reference frame buffer before being sent to the monitor. To properly apply the post filter, the decoder might require an extra buffer. The deployment will not include an increase in bit rate or a change in encoding technique. It is completely optional to apply a post filter because none of the applicable standards require it. Loop filtering, which happens inside the encoding loop, will result in deblocking. This is shown graphically in Figure 2.4.

To correct the motion in frames, initially frames are sent to filtering primarily. For the decoder to carry out filtering in the same way as the encoder, a standard conformant decoder is needed. Each CTU's output is handled independently during decoding, and A reference frame buffer holds the processed data. Both the encoding and decoding processes employ the same filtering. There's no need for additional decoder buffering. However, this complicates incorporation into commercial code packages. When Post and Loop filtering are examined side by side, their advantages and disadvantages become evident.

One critical phase in block-based video encoders is motion estimation (ME), which plays a vital role in intelligently predicting Motion Vectors (MVs) for each block in every frame. This process potentially reduces the overall bit rate of a video stream. We utilize motion estimation for this purpose. Achieving precise motion estimation within a frame leads to a higher enhancement ratio, as it generates less entropy information for the remaining frame blocks. The first thing ME needs to do is figure out which block from the newest and most recent decoded frame (the area of interest, also called the search window) most closely fits every block in the reference frame. to determine the motion content of each block in each frame. Figure 2.5 depicts the notion of ME. The ME approach finds the best feasible match by first creating a distinct Search Window (SW) based on a given cost function for every block in the present frame. The final results generated by the ME are the coordinates of the optimum MV and the corresponding cost. The Lagrangian multiplier DMV, the distortion  $D$ , the bit-rate required to encode the motion vector  $R$ , the Lagrangian cost function  $J$ , the search window,

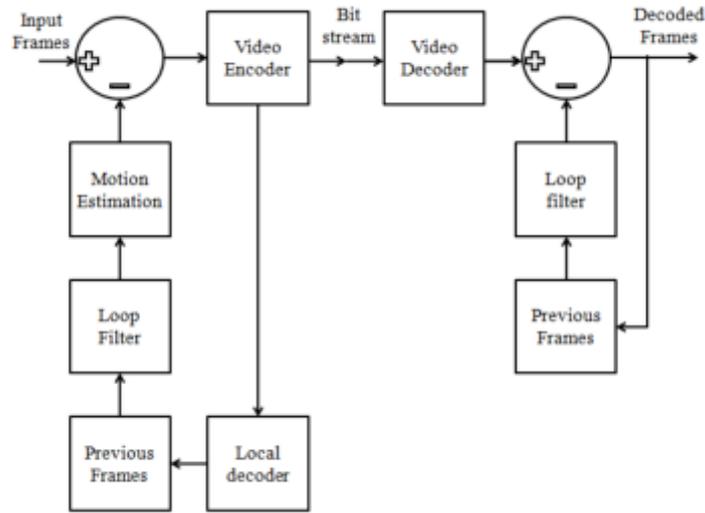


Fig. 2.4: Artifacts block reduction using loop filter

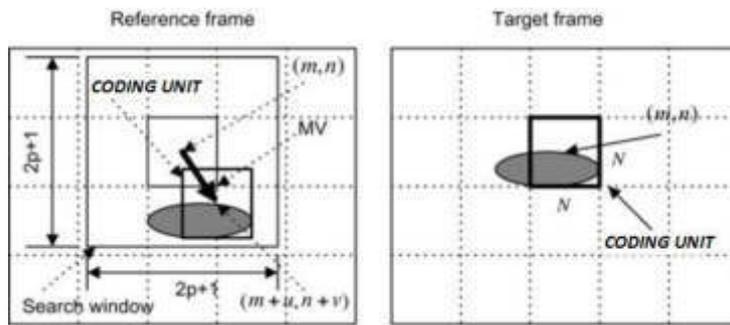


Fig. 2.5: A Process Illustration for Motion Estimation

and the optimal motion vector  $MV(x,y)$  can all be used to characterize the ME issue. Implementing an MV encoder with a bit rate of  $R$  makes it easier to solve the ME problem. The SAD or SSD distortion properties (Sum of Squared Difference) are utilized by numerous applications.

The SAD distortion feature is commonly utilized and that can

$$MV(x_i, y_j) = \min(J_{c oct}(X + i, Y_j)(X + i, Y))cSW \tag{2.1}$$

$$J_{c oct} = D + h_M V R \tag{2.2}$$

$$D = SAD(x, y) = \sum_{i=1}^M \sum_{j=1}^M |C(X_i, Y_j) - R(X_i, Y_j)| \tag{2.3}$$

We can use  $MXN$  for the block's pixel size and  $R$  for a reference block, and  $C$  for the current block to get an estimate of the size of the active block (3). Video encoders for ME that are block-based frequently use Block Matching Algorithms (BMAs). When a problem arises in the software system, the ME technique, also

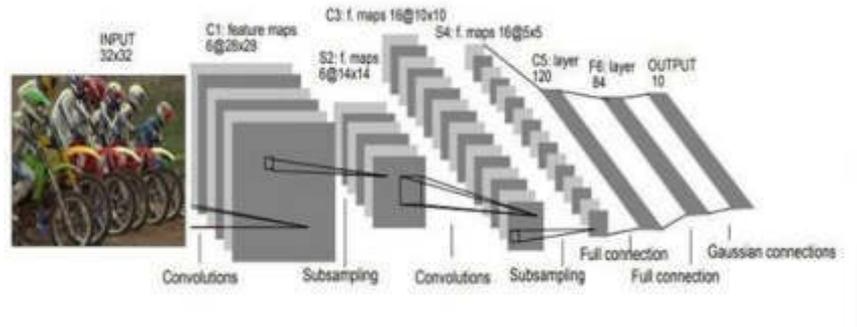


Fig. 2.6: Model of DLCNN

called the Full Search algorithm, looks into each block individually. Searching through every block in the entire software system is more challenging for the encoder. Effective ME approaches are used to bypass blocks that are less likely to have the optimal MV due to the limitations of video encoding time. Nevertheless, the final video quality (MVs) may degrade if you opt for a quick search approach instead of one that takes longer to find the best frames. Therefore, a quick and effective ME algorithm is required to reduce ME complexity while preserving an adequate ratio of enhancement and high-quality output video.

The suggested architecture for the DLCNN, illustrated in Figure 2.6, comprises convolutional blocks (CB), deconvolutional blocks (DB), attention-based shared weights blocks (ASB), attention-based shared weights residual blocks (ASRB), multi-head attention blocks (MHA), low-resolution feature fusions (LFF), and high-resolution feature fusions (HRFF). During its first stages of development, the DLCNN was inspired by three concepts: First, there's the attention mechanism that fine-tunes the feature maps using channel and spatial attention; The second idea is weight sharing, which employs the same convolutional layer twice in a single layer block; the third idea is a novel multiresolution fusion of features block design, which uses a design block to fuse current-resolution feature maps with either low- or higher-resolution feature maps.

There are three feasible patch resolutions to execute the DLCNN model: full, half, and quarter. This also includes the establishment of new hubs. The network is now better able to obtain feature maps from the input patch and forecast the specifics of the refinement at every level thanks to the addition of these additional nodes. The DLCNN helps all three patch kinds (hw, h2w2, and h4w4). To create the three feature maps, the input patch must be processed in the first stage of DLCNN. This method uses the entire patch for processing as opposed to the normal method, which only uses the input patch. In contrast to the conventional method that solely relies on full-patch resolution, the DLCNN algorithm gathers refined information from half, quarter, and full-patch resolutions. The typical approach disregards resolutions smaller than the entire patch in its standard setup. This deviation from the norm is essential.

**3. Results and Discussions.** The suggested method employed the DLCNN configuration, which produced superior visual and RD performance compared to alternative configurations. An improved visual result in random access setup is shown by the growing RD performance metrics, such as PSNR and bit rate. Additionally, resolution is used to gauge the suggested HEVC's visual analysis. The suggested technique reduced the quantity of transitional frames in the video while maintaining resolution throughout. The video's total memory usage decreases when the number of frames is decreased, but the resolution remains unchanged.

A range of video clips, such as People Street, BQTerrace, Basketball Pass and Cactus are used to assess the subjective performance of the proposed technique. These are high-definition videos, and the recommended method preserved the videos' sizes throughout. Table 3.1 displays the properties of videos in various sequences. The suggested parameters, including multiple test sequence types with varying QP and resolutions, are displayed in Table 3.1. Every aspect of human perception is dependent on rate-distortion optimization; prediction techniques, along with each of intra, random access, and low delay techniques, are crucial to this process.

The subjective assessments of the two samples' various approaches are displayed in Figures 3.2 and 3.3. In

Table 3.1: Assumption

	Suggested Values
Software Reference	HM 16.8
QP	22,27,32,37
Total Sequence	People on street, Cactus, BQ Terrance, Basketball pass
The Resolution	(2400x1600), (1920x1080), (832x480), (416x240)
Central Processing Unit	Intel Core i7 x 990, 3.47GHz
Operating System	Windows 10 (64-bits)



Fig. 3.1: A few frames from a video



Fig. 3.2: An impartial assessment of sample 1 (a) original frame, (b) HM 11[23], (c) Luo's [25], (d)PVC [34], (e) Proposed FQE-Net



Fig. 3.3: An impartial assessment of sample 2, (a) original frame, (b) HM 11[23], (c) Luo's [25], (d) PVC [34], (e) Proposed FQE-Net

this case, Proposed FQE-Net outperformed traditional methods like HM 11[23], Luo's [25], and PVC [34] in terms of subjective performance.

Table 3.2 displays the results of the block mode tests we ran. When the Motion Vector's magnitude conditions are changed, the simulated analysis produces different outcomes. Table 3.3 illustrates how content-

Table 3.2: Block mode FQE-Net results

Sequence	QP	U-PSNR	V-PSNR	Y-PSNR	Bytes written to file	YUV-PSNR	Time (sec)
PeopleStreet	23	40.78	40.90	42.90	4457095	42.52	68.489
Cactus	23	40.70	40.70	42.34	6673845	42.09	122.70
BQTerrace	23	40.40	40.40	41.72	10947093	41.61	193.779
BasketballPass	23	40.39	40.46	41.80	1164445	41.66	21.63
PeopleStreet	28	36.40	36.40	36.82	3577587	36.67	76.30
Cactus	28	36.50	36.53	37.06	5473656	36.86	118.614
BQTerrace	28	35.39	36.41	36.84	90152893	36.69	190.207
BasketballPass	28	36.39	36.44	36.86	958439	36.71	22.14
PeopleStreet	33	32.49	32.49	31.87	2921654	32.08	64.06
Cactus	33	33.59	32.59	31.91	44438764	32.11	114.741
BQTerrace	33	32.40	32.49	31.87	7354124	32.08	189.50
BasketballPass	33	31.60	32.54	31.90	782213	32.10	22.12
PeopleStreet	37	29.60	29.55	27.97	2346543	27.68	69.38
Cactus	37	29.60	29.58	27.01	35727654	27.71	114.569
BQTerrace	38	29.60	29.55	27.97	5899859	27.67	187.112
BasketballPass	38	29.60	29.56	27.98	628833	27.69	19.91

Table 3.3: Content mode results of FQE-Net

Sequence	QP	U-PSNR	V-PSNR	Y-PSNR	Bytes written to file	YUV-PSNR	Time (sec)
PeopleStreet	23	41.32	40.32	40.61	433795	40.51	15.61
Cactus	23	41.36	40.36	40.65	6585630	40.55	24.93
BQTerrace	23	41.33	41.33	41.62	10920577	41.51	40.31
BasketballPass	23	41.36	41.35	41.64	1160133	41.54	5.46
PeopleStreet	28	36.37	36.37	36.78	3572723	36.64	15.968
Cactus	28	36.40	36.40	36.82	5431977	36.67	25.067
BQTerrace	28	35.37	36.38	36.78	8998555	36.64	39.994
BasketballPass	28	36.41	36.40	36.80	9567876	36.67	5.18
PeopleStreet	33	32.51	32.51	31.84	2917686	32.05	16.068
Cactus	33	32.53	32.53	31.86	44435633	32.08	25.53
BQTerrace	33	32.51	32.51	31.84	7342523	32.05	40.42
BasketballPass	33	32.53	32.52	31.86	781365	32.07	5.212
PeopleStreet	37	29.56	29.55	26.97	2346623	27.68	15.18
Cactus	37	29.59	29.58	27.01	3572076	27.71	24.65
BQTerrace	38	29.55	29.55	26.97	5899832	27.67	40.422
BasketballPass	38	29.58	29.56	26.98	628887	27.69	5.12

based macro-blocks are more common as the threshold is lowered.

IDR data is frequently present in only the first frame of a low delay configuration. There are two HEVC variations that have much lower latency. Low-delay configuration -B is the first option that must be chosen; low-delay configuration -P is the second option that can be chosen freely. Every frame in a GOP is recorded as a P-picture in low-delay P mode; in contrast, every frame in a GOP is recorded as a Generalized P and B image (GPB) in low-delay GPB mode. This is the main distinction between the P configuration and the low-delay B configuration. In all these layouts, the original image is described using IDR encoding. It is possible to determine the QP of each altered picture by including an offset parameter.

Random Access Configuration: In a random access setup, encoding is done using a hierarchical B-structure. As can be seen in the image, the frames (designated L1 through L4) are made up of numerous layered parts.

Table 3.4: Evaluating FQE-Net's performance in a low-latency configuration

Sequence	QP	PSNR (in dB)				Bitrate (in Kbps)			
		Luo's [25]	PVC [34]	HM 11 [23]	Proposed	Luo's [25]	PVC [34]	HM 11 [23]	Proposed
BQTerrace (1920 x 1080)	23	37.85	36.7	39.24	41.06	32654	16766	52765	5454
	28	36.23	35.77	36.4	35.69	6438	5297	7523	4487
	33	34.53	34.37	34.58	30.32	1958	1934	1998	3523
	38	32.37	32.36	32.37	24.81	764	799	756	2778
Cactus (1920 x 1080)	23	39.35	38.35	39.69	40.18	16576	12465	20087	5743
	28	37.62	37.21	37.7	38.24	5687	5398	5723	4787
	33	35.55	35.47	35.56	33.75	2535	2587	2512	3812
	38	33.25	33.25	33.26	28.52	1287	1223	1298	2898
ParkScene (832 x 420)	23	40.53	39.14	40.91	43.16	7432	6287	7923	5721
	28	37.92	37.43	37.31	38.22	3165	2923	3109	4787
	33	35.18	35.13	35.18	32.82	1365	1365	1323	3799
	38	35.59	32.59	32.59	28.48	565	579	587	3034
BQMall (832 x 420)	23	40.32	38.61	41.12	43.24	3887	3223	4221	5776
	28	37.99	37.25	38.25	37.40	1854	1798	1876	4612
	33	35.24	35.09	35.28	33.00	799	823	912	3898
	38	32.28	32.37	32.38	28.54	465	498	498	3023

Table 3.5: Evaluating FQE-Net's performance in a low-latency configuration

Sequence	QP	PSNR (in dB)				Bitrate (in Kbps)			
		HM 11 [23]	Luo's [25]	PVC [34]	Proposed	MM 11 [23]	Luo's [25]	PVC [34]	Proposed
BQTerrace (1920 x 1080)	23	37.63	37.53	36.19	42.86	30360	25419	14440	6420
	28	36.19	36.12	35.45	37.84	6623	5910	5240	5170
	32	34.81	34.72	34.58	33.30	2255	2220	2210	4230
	38	32.98	32.89	32.94	28.87	970	970	980	3776
Cactus (1920 x 1080)	23	39.17	39.13	37.77	42.09	15967	14580	11420	12687
	28	37.76	37.66	37.01	37.73	5650	5390	5310	10523
	33	35.94	35.84	35.75	33.40	2682	2644	2650	8790
	38	33.77	33.71	33.75	28.29	1380	1389	1370	7146
ParkScene (832 x 420)	24	40.57	40.54	38.51	43.28	7380	6988	6087	7800
	29	38.42	39.28	37.48	38.28	3310	3145	3110	4100
	34	35.93	36.82	35.77	33.27	1540	1533	1544	2100
	39	33.44	33.38	33.44	28.26	720	718	730	1100
BQMall (832 x 420)	22	39.50	39.50	37.15	42.26	3560	3302	2928	5562
	27	37.42	37.29	36.19	36.47	1691	1621	1596	4565
	32	34.87	34.63	34.60	32.82	860	850	848	3641
	37	32.16	32.08	32.14	31.95	460	450	449	2650
PartyScene (832 x 420)	22	36.77	36.67	33.20	38.00	6638	6011	4856	4760
	27	34.25	33.69	32.57	36.60	3071	2882	2774	4304
	32	31.29	31.26	31.25	26.91	1461	1443	1439	3597
	37	28.65	28.51	28.62	24.45	692	690	690	2765
RaceHorses (832 x 420)	22	38.15	38.09	35.70	38.71	4464	4135	3240	5457
	27	35.63	35.49	34.59	32.94	1996	1894	1788	4563
	32	32.94	32.82	32.81	27.13	945	935	935	3641
	37	30.33	30.26	30.26	28.90	463	462	462	2747

An image of the IDR coding decoded is shown in the first image. The next image in our introductory image series is encoded as a GPB picture, which means it can connect to other GPB pictures as well as I-frames (for Inter-Prediction). B-grade photos are tucked away in the depth map of the picture. "B-pictures," or the lowest layer of images, are those that don't make any references to other images. This indicates that no other frame uses them as a standard. The comprehensive quantitative analysis of the bit rate (in Kbps) and PSNR (in dB) metrics for the suggested approaches utilizing Random-Access and Low delay configurations is provided in Tables 3.4 and 3.5. Ultimately, the comparative analysis between the two configurations demonstrates that the suggested approach outperforms the HM 11[23], Luo's [25], and perceptual video coding PVC [34] in terms of rate distortion performance.

It appears that you have shared details regarding a structure that improves the efficiency of video compression and transmission. Please let me know if you need help reframing or extending this information. This is an updated version: "When compared to modern encoders, the architecture described provides better video compression while using less bandwidth. Its low latency and compatibility with 5G network transmission speed distinguish it as a state-of-the-art solution. The H.265 protocol architecture's incorporation of Low Entropy (LE) is essential for decreasing buffering latency and raising overall efficiency. Apart from these enhancements, the quality of video streaming is given special attention in this study. To measure and quantify the video quality, evaluation measures such video multi-method assessment fusion (VMAF), peak signal-to-noise ratio (PSNR), and structural similarity index (SSIM) were used. A thorough analysis comparing latency and compression ratios for designs with and without H.264 shows that the suggested architecture performs noticeably better. In summary, deep learning is a potential technique for cutting-edge video streaming systems because it not only increases accuracy but also lowers computational complexity when used for efficient data transmission. By producing an accurate fused prediction block, inter bi-prediction is a vital tool in the field of video coding, greatly increasing coding efficiency. Even with the incorporation of block-wise techniques such as bi-prediction with CU-level weight (BCW) into Versatile Video Coding (VVC), linear fusion-based schemes continue to face difficulties in accurately portraying a range of pixel fluctuations inside a block. Bi-directional optical flow (BDOF), a pixel-wise method, has been devised to improve bi-prediction blocks in order to overcome these drawbacks. Nevertheless, the BDOF mode's non-linear optical flow equation functions based on assumptions, which limits its ability to precisely adjust for different kinds of bi-prediction blocks.

**4. Conclusion.** This work primarily aims to implement a deep learning-based network for enhancing quality on a frame-by-frame basis. With the energy compaction, FQE-Net, artifacts, complexity, and efficiency are all being reduced, allowing the HEVC codec to support videos up to 16K in resolution. The DLCNN model must be used in place of the HEVC codec's DCT and DWT transforms in order to increase efficiency. Along with the Content Block Search Algorithm, we will also employ filtering techniques like the Sample Adoptive Filter and the Deblocking Filter for Motion Estimation and Compensation. The simulation's results indicate that the suggested FQE-Net appears to be more efficient than earlier techniques.

#### REFERENCES

- [1] Zhang, Zhenzhen, et al. "A CNN-Based HEVC Video Steganalysis Against DCT/DST-Based Steganography." International Conference on Digital Forensics and Cyber Crime. Springer, Cham, 2022.
- [2] Kuanar, Shiba, et al. "Deep learning based HEVC in-loop filter and noise reduction." *Signal Processing: Image Communication* 99 (2021): 116409.
- [3] Dhanalakshmi, A., & Nagarajan, G. (2022). Group-normalized deep CNN-based in-loop filter for HEVC scalable extension. *Signal, Image and Video Processing*, 16(2), 437-445.
- [4] Liu, J., Li, Z., Jiang, X., & Zhang, Z. (2021). A high-performance CNN-applied HEVC steganography based on diamond-coded PU partition modes. *IEEE Transactions on Multimedia*, 24, 2084-2097.
- [5] Zhu, Chen, Yan Huang, Rong Xie, and Li Song. "HEVC VMAF-oriented Perceptual Rate Distortion Optimization using CNN." In 2021 Picture Coding Symposium (PCS), pp. 1-5. IEEE, 2021.
- [6] Issa, O., & Shanableh, T. (2022). CNN and HEVC Video Coding Features for Static Video Summarization. *IEEE Access*, 10, 72080-72091.
- [7] Bouaafia, S., Khemiri, R., Messaoud, S., & Sayadi, F. E. (2022). Deep CNN Co-design for HEVC CU Partition Prediction on FPGA-SoC. *Neural Processing Letters*, 1-19.
- [8] Amna, Maraoui, Werda Imen, Bouaafia Soulef, and Sayadi Fatma Ezahra. "Machine Learning-Based approaches to reduce HEVC intra coding unit partition decision complexity." *Multimedia Tools and Applications* 81, no. 2 (2022): 2777-2802.

- [9] Zaki, F., Mohamed, A. E., & Sayed, S. G. (2021). CtuNet: a deep learning-based framework for fast CTU partitioning of H265/HEVC intra-coding. *Ain Shams Engineering Journal*, 12(2), 1859-1866.
- [10] Li, Z., Jiang, X., Dong, Y., Meng, L., & Sun, T. (2022). An anti-steganalysis HEVC video steganography with high performance based on CNN and PU partition modes. *IEEE Transactions on Dependable and Secure Computing*.
- [11] Bouaafia, S., Khemiri, R., Maraoui, A., & Sayadi, F. E. (2021). CNN-LSTM learning approach-based complexity reduction for high-efficiency video coding standard. *Scientific Programming*, 2021.
- [12] Imen, Werda, Maraoui Amna, Belghith Fatma, Sayadi Fatma Ezahra, and Nouri Masmoudi. "Fast HEVC intra-CU decision partition algorithm with modified LeNet-5 and AlexNet." *Signal, Image and Video Processing* (2022): 1-9.
- [13] Liu, C., Jia, K., & Liu, P. (2021). Fast Depth Intra Coding Based on Depth Edge Classification Network in 3D-HEVC. *IEEE Transactions on Broadcasting*, 68(1), 97-109.
- [14] Huang, Yan, Li Song, Rong Xie, Ebroul Izquierdo, and Wenjun Zhang. "Modeling acceleration properties for flexible INTRA HEVC complexity control." *IEEE Transactions on Circuits and Systems for Video Technology* 31, no. 11 (2021): 4454-4469.
- [15] Yang, R., Zhu, S., Zheng, X., & Zeng, B. (2022). OL-DN: Online learning-based dual-domain network for HEVC intra-frame quality enhancement. *arXiv preprint arXiv:2208.04661*.
- [16] Dai, Y., Xue, C., & Zhou, L. (2022). Visual saliency guided perceptual adaptive quantization based on HEVC intra-coding for planetary images. *Plos one*, 17(2), e0263729.
- [17] Ding, Q., Shen, L., Yu, L., Yang, H., & Xu, M. (2021). Patch-wise spatial-temporal quality enhancement for HEVC compressed video. *IEEE Transactions on Image Processing*, 30, 6459-6472.
- [18] Yang, Renwei, Hwei Liu, Shuyuan Zhu, Xiaozhen Zheng, and Bing Zeng. "DFCE: Decoder-Friendly Chrominance Enhancement for HEVC Intra Coding." *IEEE Transactions on Circuits and Systems for Video Technology* (2022).
- [19] Pan, Zhaoqing, Weijie Yu, Jianjun Lei, Nam Ling, and Sam Kwong. "TSAN: Synthesized view quality enhancement via two-stream attention network for 3D-HEVC." *IEEE Transactions on Circuits and Systems for Video Technology* 32, no. 1 (2021): 345-358.
- [20] Yue, Jian, Yanbo Gao, Shuai Li, Hui Yuan, and Frédéric Dufaux. "A Global Appearance and Local Coding Distortion Based Fusion Framework for CNN Based Filtering in Video Coding." *IEEE Transactions on Broadcasting* 68, no. 2 (2022): 370-382.
- [21] Pan, Z., Yuan, F., Yu, W., Lei, J., Ling, N., & Kwong, S. (2022). RDen: Residual distillation enhanced network-guided lightweight synthesized view quality enhancement for 3D-HEVC. *IEEE Transactions on Circuits and Systems for Video Technology*.
- [22] Galiano, V., Migallón, H., Martínez-Rach, M., López-Granado, O., & P Malumbres, M. (2022). On the use of deep learning and parallelism techniques to significantly reduce the HEVC intra-coding time. *The Journal of Supercomputing*, 1-19.
- [23] Wang, Zixi, and Fan Li. "Convolutional neural network based low complexity HEVC intra encoder." *Multimedia Tools and Applications* 80.2 (2021): 2441-2460.
- [24] Sheeba, G., & Maheswari, M. (2022). HEVC Video Quality Enhancement Using Deep Learning with Super Interpolation and Laplacian Filter. *IETE Journal of Research*, 1-14.
- [25] Dong, Yi, Xinghao Jiang, Zhaohong Li, Tanfeng Sun, and Zhenzhen Zhang. "Multi-Channel HEVC Steganography by Minimizing IPM Steganographic Distortions." *IEEE Transactions on Multimedia* (2022).
- [26] Jayaraman, Thiyagarajan, and Gowri Shankar Chinnusamy. "Analysis of deep rain streaks removal convolutional neural network-based post-processing techniques in HEVC encoder." *Journal of Circuits, Systems and Computers* 30.02 (2021): 2150020.
- [27] Pan, Z., Zhang, P., Peng, B., Ling, N., & Lei, J. (2021). A CNN-based fast inter coding method for VVC. *IEEE Signal Processing Letters*, 28, 1260-1264.
- [28] Li, J., Yang, M., Xie, Y., & Li, Z. (2022). A CU Depth Prediction Model Based on Pre-trained Convolutional Neural Network for HEVC Intra Encoding Complexity Reduction. In *Advances in Machine Learning/Deep Learning-based Technologies* (pp. 217-231). Springer, Cham.
- [29] Uddin, K., Yang, Y., & Oh, B. T. (2022). Double compression detection in HEVC-coded video with the same coding parameters using picture partitioning information. *Signal Processing: Image Communication*, 103, 116638.
- [30] Ding, D., Kong, L., Wang, W., & Zhu, F. (2021). A progressive CNN in-loop filtering approach for inter frame coding. *Signal Processing: Image Communication*, 94, 116201.
- [31] Shao, T., Liu, T., Wu, D., Tsai, C. Y., Lei, Z., & Katsavounidis, I. (2022). PTR-CNN for in-loop filtering in video coding. *Journal of Visual Communication and Image Representation*, 103615.
- [32] Yao, Haichao, Rongrong Ni, Hadi Amirpour, Christian Timmerer, and Yao Zhao. "Detection and Localization of Video Transcoding From AVC to HEVC Based on Deep Representations of Decoded Frames and PU Maps." *IEEE Transactions on Multimedia* (2022).
- [33] Bouaafia, Soulef, Randa Khemiri, Seifeddine Messaoud, Olfa Ben Ahmed, and Fatma Ezahra Sayadi. "Deep learning-based video quality enhancement for the new versatile video coding." *Neural Computing and Applications* 34, no. 17 (2022): 14135-14149.
- [34] Balaji, L., A. Dhanalakshmi, and C. Raja. "Scalable extension of HEVC with adaptive de-blocking filter for bandwidth-limited applications." *ICT Express* (2021).
- [35] Esmaeeli, H., & Rezaei, M. (2022). A content-based intra rate-distortion model for HEVC-SCC. *Multimedia Tools and Applications*, 81(12), 16515-16536.
- [36] Zhao, J., Dai, P., & Zhang, Q. (2021). A Complexity Reduction Method for VVC Intra Prediction Based on Statistical Analysis and SAE-CNN. *Electronics*, 10(24), 3112.
- [37] Nami, S., Pakdaman, F., Hashemi, M. R., & Shirmohammadi, S. (2022). BL-JUNIPER: A CNN-Assisted Framework for

- Perceptual Video Coding Leveraging Block-Level JND. IEEE Transactions on Multimedia.
- [38] Jia, M., Gao, Y., Li, S., Yue, J., & Ye, M. (2021). An explicit self-attention-based multimodality CNN in-loop filter for versatile video coding. *Multimedia Tools and Applications*, 1-15.
- [39] Feng, C., Qi, Z., Danier, D., Zhang, F., Xu, X., Liu, S., & Bull, D. (2022). Enhancing HDR Video Compression through CNN-based Effective Bit Depth Adaptation. *arXiv preprint arXiv:2207.08634*.
- [40] Ravishankar, Joshitha, Mansi Sharma, and Pradeep Gopalakrishnan. "A flexible coding scheme based on block krylov subspace approximation for light field displays with stacked multiplicative layers." *Sensors* 21, no. 13 (2021): 4574.

*Edited by:* Anil Kumar Budati

*Special issue on:* Soft Computing and Artificial Intelligence for wire/wireless Human-Machine Interface

*Received:* Oct 4, 2023

*Accepted:* Dec 29, 2023



## SOFTWARE EFFORT ESTIMATION USING MACHINE LEARNING ALGORITHMS

KRUTI LAVINGIA <sup>\*</sup>, RAJ PATEL <sup>†</sup>, VIVEK PATEL <sup>‡</sup> AND AMI LAVINGIA <sup>§</sup>

**Abstract.** Effort estimation is a crucial aspect of software development, as it helps project managers plan, control, and schedule the development of software systems. This research study compares various machine learning techniques for estimating effort in software development, focusing on the most widely used and recent methods. The paper begins by highlighting the significance of effort estimation and its associated difficulties. It then presents a comprehensive overview of the different categories of effort estimation techniques, including algorithmic, model-based, and expert-based methods. The study concludes by comparing methods for a given software development project. Random Forest Regression algorithm performs well on the given dataset tested along with various Regression algorithms, including Support Vector, Linear, and Decision Tree Regression. Additionally, the research identifies areas for future investigation in software effort estimation, including the requirement for more accurate and reliable methods and the need to address the inherent complexity and uncertainty in software development projects. This paper provides a comprehensive examination of the current state-of-the-art in software effort estimation, serving as a resource for researchers in the field of software engineering.

**Key words:** Software Engineering, Machine Learning, Effort Estimation

**1. Introduction.** A large volume of data is produced while software companies develop and generate software [1]. From the requirements phase until maintenance, a unique collection of data is generated at each stage of software development. In software development project management, key factors such as Lines of Code (LOC), historical project data, team skill levels, team size, functional and non-functional requirements, project phases, and development timelines are crucial for success. LOC quantifies code changes, while past project data informs resource allocation. Team skills and size impact productivity, while requirements shape the project's direction [2]. Monitoring project phases and adhering to timelines ensures progress and identifies bottlenecks. These elements collectively enable effective project management, guiding teams to deliver software solutions that align with stakeholder expectations and meet project goals efficiently.

The data produced in software repositories is collected and maintained by software organizations as part of their ongoing efforts to improve the software quality. Several data mining techniques are used to analyze the vast amounts of data kept in software repositories to uncover new patterns and standout data points [3].

Two-thirds of all large projects greatly exceed their original projections and one-third of projects go over budget and are delivered late, as claimed by surveys [4].

The technique of estimating the effort and resources needed to construct a software system is known as software engineering cost estimation. Managers and stakeholders typically use this process to plan and budget software development projects. Cost estimation methods can range from simple rule-of-thumb calculations to more formal methods, such as parametric modelling or expert judgment. Factors that can influence the cost of a software project include the size and complexity of the system, the development methodologies and tools used, and the skill level of the development team [5]. It is essential to note that cost estimation is an iterative process that needs to be repeatedly refined and updated as more information becomes available throughout the project.

In software engineering, effort estimation is the process of predicting the number of human resources, measured in person-hours or person-months, needed to complete any software development project. It is a critical aspect of project management as it helps stakeholders to plan and budget for the project and to make

---

<sup>\*</sup>Nirma University ([kruti.lavingia@nirmauni.ac.in](mailto:kruti.lavingia@nirmauni.ac.in))

<sup>†</sup>Institute of technology, Nirma University ([20bce218@nirmauni.ac.in](mailto:20bce218@nirmauni.ac.in))

<sup>‡</sup>Institute of technology, Nirma University ([20bce226@nirmauni.ac.in](mailto:20bce226@nirmauni.ac.in))

<sup>§</sup>Sal College of Engineering ([ami.lavingia@sal.edu.in](mailto:ami.lavingia@sal.edu.in))

informed decisions about resource allocation and project scope [6].

Software engineering has a variety of effort estimating techniques, each having advantages and disadvantages. Among the more popular techniques are:

### 1.1. Methods for effort estimation.

- **Expert judgment:** This method relies on the experience and expertise of individuals who have previously worked on similar projects. It is often used as a starting point for effort estimation and can provide a rough estimate of the required effort. However, it is subject to bias and can be affected by an individual's experience [7].
- **Analogous estimation:** This method uses the data from previous similar projects to estimate the effort required for the current project. It is a quick and easy method, but it is not always accurate, as the projects may not be entirely similar [8].
- **Three-point estimation:** This method uses the most likely, optimistic, and pessimistic estimates of effort to generate a range of possible values. It helps generate a range of likely effort estimates, but it is a relatively complex method [9].
- **Parametric estimation:** This method uses mathematical models to estimate the effort required for a project. It is based on the project's size, complexity, and other [10].

Estimating is crucial in project management, as inaccuracies in estimation can lead to poor project performance, potentially resulting in project failure. One of the management factors that cause about 65% of lost projects is poor estimation technique [11]. This study uses machine learning to create a model for software cost estimation. As a result, this review aims to test if machine learning is a better technique than using traditional methods to estimate software development effort or vice versa. Support machine learning algorithms such as vector regression, regression algorithms such as simple linear regression, and decision tree-based regression are applied in this study with the assistance of the Python programming language.

The remainder of this paper is broken up into related work that focuses on previous studies that have been done in this particular area. The following section examines machine learning methods and software cost estimation. They utilized machine-learning techniques, data sets, and evaluation standards are thoroughly detailed in the next section. The comparison and in-depth analysis of the experimental results come before the conclusion and section on future work.

However, we do wish to stress the purpose of this paper is to consider how different prediction systems perform under the same conditions and how to evaluate them, not to argue in favor of any particular prediction technique.

**2. Related work.** Numerous research has suggested various models for calculating the cost of the software. To find alternatives, improve upon, or support existing models, multiple models have been proposed and constructed [12]. Various new models have been developed and constructed to discover alternatives, improve existing models, or assist current models. A well-known method for estimating software costs is the build cost model.

**2.1. Without Machine Learning.** As introduced by Barry Boehm, the Constructive Cost Model (COCOMO) stands out as the predominant approach within the algorithmic methods category [13]. It relies on a series of equations and parameters derived from past software project experiences for estimation, and its models have garnered widespread practical acceptance. In the context of COCOMO, code size is measured in Thousand Lines Of Code (KLOC), and effort is expressed in person-months. COCOMO is a valuable tool to gauge the quality and effort required for software projects, as exemplified by its application in Manikavelan's study [14], providing approximate estimates within fixed time frames. Moreover, the authors in this particular research extended COCOMO's capabilities by incorporating the Gaussian Membership Function, revealing outstanding performance of the fuzzy-COCOMO model in terms of reducing relative errors.

In the study conducted by Nandan and Deepak [15], a novel approach was employed. They utilized a hybrid BATGSA algorithm to optimize the COCOMO model, drawing data from NASA databases. The study comprehensively compared three distinct techniques implemented using MATLAB. The outcome was a noteworthy decrease in normalized error with the updated COCOMO model.

Notably, the authors introduced an innovative hybrid strategy that amalgamated fuzzy clustering, ABE,

and ANN approaches to enhance the accuracy of effort estimation. This novel approach entailed clustering all projects within a newly established framework, effectively mitigating the influence of inconsistent and irrelevant projects on projections. This research resulted in significant improvements, with an average enhancement of 0.25 in the first dataset and remarkable gains of 52 and 94 per cent in the second dataset, as demonstrated by the prediction percentage (PRED) and mean magnitude of relative error (MMRE) performance indicators.

**2.2. With Machine Learning.** In their study, Shukla et al. leveraged the Desharnais dataset to explore the performance of various machine learning models in estimating software project effort [16]. Notably, their MLPNN model achieved an R2 value of 0.79380, surpassing other models like LR, SVM, and KNN. It effectively explained 79% of the estimated variance, with only marginal differences (6-7%) in R2 values among them.

The research delved into the association between the most correlated elements by Pearson correlation and the effort variable using seven machine learning methods, following an initial correlation analysis of each dataset variable with the effort variable. Performance evaluation was based on error values [17].

Sarro introduced an effort prediction technique combining Confidence Interval Analysis and Mean Absolute Error assessment [18]. This innovative approach demonstrated promise through trials involving over 700 software programs, finding applications in diverse fields like pharmaceutical research, biochemistry, and computer vision. The method selected feature subsets based on optimization techniques and transferred them to classifiers (SVM, ANN, and Decision Tree) for classification and regression tasks involving two optimization algorithms and three classifiers. This process, known as Feature Selection, yielded excellent results across various datasets.

In another study, 93 projects' preprocessed COCOMO NASA benchmark data were employed to make predictions using machine learning techniques like Naive Bayes, Logistic Regression, and Random Forests [19]. Performance evaluation metrics such as Classification Accuracy, Precision, Recall, and AUC were employed following five-fold cross-validation. Each method outperformed the benchmark COCOMO model in production prediction.

V. Anandhiin's investigation focused on regression techniques, notably the M5 algorithm and Linear Regression, for estimating software cost using the Constructive Cost Model dataset [20]. The results indicated that the M5 method exhibited more minor errors, including the mean magnitude of the relative error and Median magnitude of the relative error, compared to Linear Regression in prediction. These clear distinctions highlight where different methods are introduced and provide insights into the authors' approach and findings.

**3. Machine Learning.** In this section, methods and ML algorithms are discussed; after that information about the dataset is stated along with its structure. Finally, the evaluation standards are the topic of conversation. Below stated section provides a clear and concise summary of some machine learning algorithms that are used to predict the effort for the project.

Machine learning techniques are increasingly thought to be crucial in research. The results of ML approaches are consistently reliable, and they are frequently employed reliable in numerous studies. Using two machine learning methods, Yeha and Deng provided a system to forecast the software product life cycle [21]. The study provided a more accurate and adaptable model for estimating product costs.

To distinguish between different types of breast cancers, Aleriza and Bitu, used support vector machines, K-nearest neighbors, and neural network classifiers [22].

Other studies concentrated more on the environment for cost assessment and other relevant elements, such as the software development life cycle relevant to the particular project. For instance, in 2018, research on the impact of organizational factors were published by Rahikkala et al., which looked at how its many components could potentially influence and improve the software cost planning process [23].

**3.1. Linear Regression.** Based on other attributes' values, linear regression analysis will predict a variable's value [24]. Two types of variables are there in the algorithm, one dependent and another independent. The dependent variable is predicted by an algorithm. The dependent variable is predicted using the independent variable as a basis. Such analysis determines the coefficients of linear mathematical equations using independent variables that may most effectively anticipate the value of the dependent variable. This algorithm fits the output on a straight line to reduce the discrepancy between the actual and anticipated output. The value of A (the dependent variable/attribute) is then estimated from B (the independent attribute/variable).

The technique used is straightforward and comparative. It is less complex than other methods for predictions. The equation for linear regression:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + e \quad (3.1)$$

$Y$  is the dependent variable and  $\beta_1, \beta_2, \dots, \beta_n$  are coefficients and  $X_1, X_2, \dots, X_n$  are the independent variables. Here  $\beta_0$  is the intercept of the line which is generated by the vertical axis.  $e$  is an error term; it is the random error used to express some random factors' effect of some random factors on  $Y$  (dependent variables).

**3.2. Support Vector Regression.** Drucker et al. initially presented the Support Vector Regression (SVR) model for regression analysis in 1997 [25].

Convex quadratic programming issues are replaced with convex linear system solutions in the least-squares SVR (LS-SVR) [26] [27], which greatly speeds up training. An extensive empirical investigation has revealed that the LS-generalization SVR's performance is comparable to the SVR's. (Van Gestel et al., 2004) [26].

Among these, Vapnik's SVR, which employs regression analysis via the support vector machine (SVM), has a wide range of applications in the energy-prediction industries. The principle of structural risk minimization underpins SVR. It has clear advantages for small datasets and can maintain excellent generalization ability [28].

In support vector regression, there are two types of hyperparameters, first being the *Kernel* function and second is a variable  $C$  which is defined as the penalty parameter of the error term. Kernel parameters affect the separation boundary. Since SVR is a kernel-based method, the performance of SVR heavily depends on the kernel functions. Different kernel equations are there, which can be applied to get the different decision boundaries.

These kernels project the input data into several high-dimensional feature spaces. Because high-dimensional feature spaces perform so well, creating new kernel functions can surpass SVR performance thresholds.

The distribution and shape of the dataset affect how the kernel is used. Here, the objective function for SVR's normal vector's size is  $|w|$ , which is minimized [26].

$$\mathbf{x}_i \mathbf{w} + \mathbf{b} = \mathbf{0} \quad (3.2)$$

$$\mathbf{x}_i \mathbf{w} + b \geq +1, y_i = +1 \quad (3.3)$$

$$\mathbf{x}_i \mathbf{w} + b \leq -1, y_i = -1 \quad (3.4)$$

$$y_i(\mathbf{x}_i \mathbf{w} + b) - 1 \geq 0, \quad \forall i \quad (3.5)$$

$$\text{minimize } \frac{\|\mathbf{w}\|^2}{2} \quad (3.6)$$

$$\text{Maximize } W(\alpha) = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j=1}^l \alpha_i \alpha_j y_i y_j (\mathbf{x}_i \mathbf{x}_j) \quad (3.7)$$

**3.3. Random Forest Regression.** Tin Kam Ho first introduced random forests in 1995. For the creation of a decision tree, the Random forests method is applied, which solves problems of classification and regression.

Later, Breiman extended the technique by combining bagging with features selected at random [29], allowing for the controlled construction of multiple decision trees using variance. Compared to decision trees, the random forest algorithm provides more accurate error rate estimates. More particular, it has been demonstrated mathematically that the error rate always tends to converge as the number of trees rises. In general, because they can quickly adjust to nonlinearities detected in the data, random decision forests tend to predict better than linear regression [30]. Random forest regression produces better results compared to other algorithms, such as support vector machines and Neural Networks, and it is also robust against over fitting [31]. This algorithm can forecast the result by running an unpruned regression on each n-ary tree from the training data and then combining the results of the n-ary tree forecasts.

**3.4. Decision Tree Regression.** A tree-based structure called decision tree regression is used to forecast the dependent variable's numerical results. Quinlan's M5 algorithm [32] is implemented using what is known as the M5P algorithm. M5P is a tree-based structure similar to CART (Classification And Regression Trees). M5P-based trees have a multivariate linear model, whereas regression trees had values at the leaves. The trees generated by the classification and regression trees are generally more prominent than M5P-generated model trees.

A tree is built using the usual decision-tree method. This decision tree employs splitting criteria to account for intra-subset variation in the class values of the samples that go down each branch. The formula below can be used to calculate the standard deviation decrease.

$$SDR = sd(T) - \sum_i \frac{|T_i|}{|T|} sd(T_i) \quad (3.8)$$

The abrupt discontinuities that would inevitably arise between nearby linear models at the pruned tree's leaves are corrected by using a method.

After the machine learning algorithms had been applied to the necessary datasets, five key statistical indicators were used as performance and assessment criteria to evaluate the success of the algorithms. Indices [33] are Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Relative Absolute Error (RAE), Root Relative Squared Error (RRSE), and Correlation Coefficient  $R^2$ .

In essence, they are utilizing the ML algorithm to calculate the error between predicted effort and actual effort found in the dataset. Assuming that  $A$  is the real effort (dependent variable, to be predicted),  $\bar{A}$  is the mean of  $A$ , and  $n$  is the number of individual data points available. The following formulae can be used to compute the error measures. Equations for the indices are:

$$MAE = \frac{1}{n} \sum_{i=1}^n |A_i - \tilde{A}_i| \quad (3.9)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (A_i - \tilde{A}_i)^2} \quad (3.10)$$

$$RAE = \frac{\sum_{i=1}^n |A_i - \tilde{A}_i|}{\sum_{n=1}^n |A_i - \bar{A}_i|} \quad (3.11)$$

$$RRAE = \sqrt{\frac{\sum_{i=1}^n |A_i - \tilde{A}_i|}{\sum_{n=1}^n |A_i - \bar{A}_i|}} \quad (3.12)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (A_i - \tilde{A}_i)^2}{\sum_{n=1}^n (A_i - \bar{A}_i)^2} \quad (3.13)$$

## 4. Research Methodology.

**4.1. Data Pre Processing.** The data utilized in this work is from a NASA software engineering dataset (NASA93) that is publicly available and contains information on several projects that have undergone development at NASA throughout the years. The dataset is structured in ARFF (Attribute Relation File Format) format and comprises 93 rows and 26 columns. The nominal values have been transformed into comparable values for better training. After the descriptive columns that didn't help with training were removed, they were discarded. To make the dataset easier to work with throughout the computation, it is then transformed into the comma-separated values format. Table 4.1 shows a detailed description of the data being used to train the various models

Table 4.1: Description of the data

Attribute	Symbol	Description	Datatype
Project	project name	Name	String
Category of application	cat2	Which field is this project related to	String
System	forg	Flight system or Ground system	character [ f , g ]
NASA center	center	Which NASA center had worked on the project	Number between 1 to 6
Capability of analyst	acap	Increase these to decrease effort	Positive integer
Capability of programmers	pcap		
Domain experience	aexp		
Current programming techniques	modp		
Software tool usage	tool		
Experience with programming languages	lexp		
Experience with VM	vexp		
Time restriction	sced		
The primary memory restriction	stor	Increase these to decrease effort	Positive integer
Database size	data		
Runtime restriction on CPU	time		
Turnaround time	turn		
Machine volatility	virt		
The difficulty of the process	cplx		
Required Software reliability	rely		
Equivalent physical line of code	equivphyskloc	Kilo lines of code	Positive integer
Development effort	act_effort	The effort in terms of person-month	Positive integer

Table 4.2: Parameters By GridSearchCV for SVR

Variable	Value
c	10
Gamma	Auto
Kernel	Linear

**4.2. Model Training.** “Development effort” is the primary dependent variable that is under study. The dataset’s actual value will be compared to the expected value as part of the prediction process carried out by the Machine Learning algorithms. The given dataset predominantly consists of more than 70% numerical data. The data set is then divided at random into a training set and a testing set at a ratio of 70:30.

**4.3. Evaluation, Results, and Discussion.** In this section, the result of the experiment is discussed and displayed. The default parameters were employed for the LinearRegression and DecisionTreeRegression models. The parameters for SupportVectorRegression are displayed in the table 4.2. In the case of RandomForestRegression, all default parameters are utilized except for the specification that sets *max\_depth* to 4. In table 4.2, Best parameters returned by GridSearchCV algorithm for SVR is shown.

Figures 4.1 to 4.3 shows the relationship between various independent variable and dependent variable.

Figure 4.4 provides a clear illustration of the relationship between CPU runtime restrictions and both database size and memory requirements. As the database size and memory requirements expand, the CPU

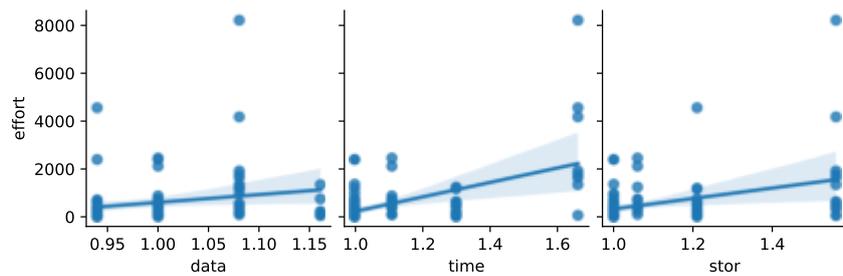


Fig. 4.1: Plotting Dependent Variable(Effort) vs. Independent Variables

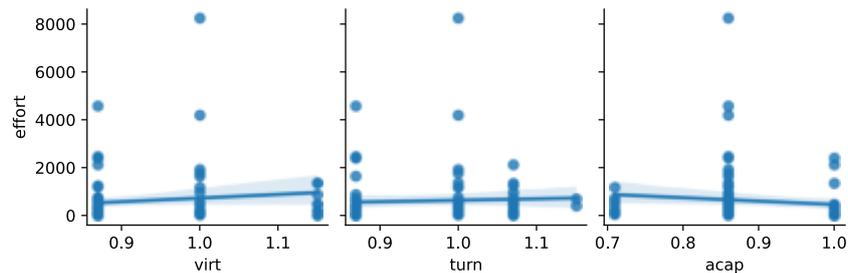


Fig. 4.2: Plotting Dependent Variable(Effort) vs. Independent Variables

experiences increased runtime restrictions due to the necessity for it to handle a greater volume of tasks within the same time frame. Furthermore, an interesting correlation emerges: an escalation in CPU runtime restrictions corresponds to a concurrent increase in the required software reliability. This suggests a dependency between these variables.

Additionally, it is noteworthy that as an analyst's capability improves, there is a parallel enhancement in their domain familiarity. This underscores the positive relationship between analyst competency and domain knowledge acquisition.

The outcome obtained using various error-measuring methods on various models is shown in table 4.3 demonstrates that the Random Forest method's anticipated value and the actual value have a very close connection. Usually  $R^2$  values are between  $-1$  and  $1$ . For highly correlated data  $R^2$  value is closer to  $1$ . Random Forest has the highest correlation coefficient compared to all other models. The other techniques for model comparison include MAE and RAE%, RMSE and RRSE%. RMSE is commonly utilized and is regarded as a preferable all-purpose error measure for numerical forecasts. In general, random decision forests tend to predict better than linear regression because they can quickly adjust to nonlinearities detected in the data. Since the Root Mean Squared Error value for Support Vector Regression is smaller, it suggests that its predictions on the test data are more accurate compared to the other models used in this study.

**4.4. Conclusion.** To establish the cost required staff, and schedule for software development, it is the job of the project manager to estimate the effort of development. The findings of several studies indicate that early-stage project estimating errors is the primary cause of software project failures. The performance of any estimating approach depends on a number of factors, including project complexity, project duration, personnel skill, development process, and others. The usage of several cost-estimating methodologies is reviewed in this essay. This paper's contribution is the improvement of our understanding of the subject of research provided by the literature review. No approach, namely along the RMSE dimension, estimates software development effort especially well in the absolute sense but comparing relatively, Random Forest Regression yields the best results as it has the lowest  $R^2$  and MAE. Additionally, practical applications of ML-based approaches could

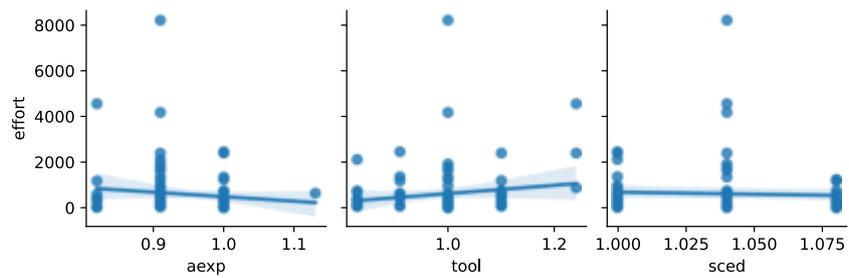


Fig. 4.3: Plotting Dependent Variable(Effort) vs. Independent Variables

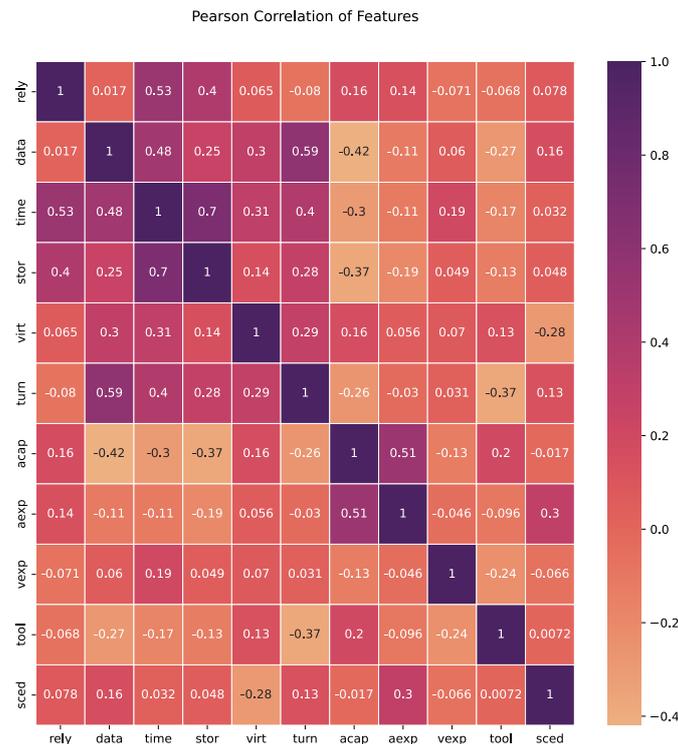


Fig. 4.4: Heat map showing Correlation between different Independent variables

involve project managers and data scientists working collaboratively to gather relevant data, including historical project data, developer expertise, project complexity, and duration. By employing machine learning algorithms on this data, project managers can enhance their ability to make more accurate early-stage project estimates, potentially reducing the risk of project failures and improving overall project management. Future research should focus on selecting optimal metrics for cost assessment, particularly leveraging computational intelligence methods.

REFERENCES

[1] Dhamija, A. & Sikka, S. A Review Paper on Software Engineering Areas Implementing Data Mining Tools & Techniques. *International Journal Of Computational Intelligence Research (IJCIR)*. **13** pp. 559-574 (2017,5)

Table 4.3: Outcomes of ML models using error measuring indices

Method Name	MAE	RMSE	RAE	RRSE	R <sup>2</sup>
Linear Regression	760.11	1365.5	1.45	1.24	-0.41
Support Vector Regression	178.7	323.61	1.05	1.042	-0.067
Random Forest Regression	642.5	1481.7	0.85	1.070	0.2781
Decision Tree Regression	760.0	1666.1	0.82	1.098	0.0873

- [2] Srinivasan, K. & Fisher, D. Machine learning approaches to estimating software development effort. *IEEE Transactions On Software Engineering*. **21**, 126-137 (1995)
- [3] Sharma, P. & Singh, J. Systematic Literature Review on Software Effort Estimation Using Machine Learning Approaches. *2017 International Conference On Next Generation Computing And Information Systems (ICNGCIS)*. pp. 43-47 (2017)
- [4] Moloekken-OEstvold, K., Joergensen, M., Tanilkan, S., Gallis, H., Lien, A. & Hove, S. A survey on software estimation in the Norwegian industry. *10th International Symposium On Software Metrics, 2004. Proceedings..* pp. 208-219 (2004)
- [5] Dakwala, A. & Lavingia, K. A Machine learning approach to improve the efficiency of Fake websites detection Techniques.
- [6] Walkerden, F. & Jeffery, R. An Empirical Study of Analogy-based Software Effort Estimation. *Empirical Software Engineering*. **4**, 135-158 (1999,6,1), <https://doi.org/10.1023/A:1009872202035>
- [7] Hughes, R. Expert judgement as an estimating method. *Information And Software Technology*. **38**, 67-75 (1996), <https://www.sciencedirect.com/science/article/pii/0950584995010459>
- [8] Shepperd, M. & Schofield, C. Estimating software project effort using analogies. *IEEE Transactions On Software Engineering*. **23**, 736-743 (1997)
- [9] Royce, W. Managing the development of large software systems: concepts and techniques. *Proceedings Of The 9th International Conference On Software Engineering*. pp. 328-338 (1987)
- [10] Baker, F. & Kim, S. Item response theory: Parameter estimation techniques. (CRC press,2004)
- [11] McManus12, J. & Wood-Harper, T. Understanding the sources of information systems project failure. (2007)
- [12] Ami, R., Mehta, V. & Lavingia, K. Analyzing the non-linear effects in DWDM optical network using MDRZ modulation format. *International Journal Of Advance Engineering And Research Development (IJAERD) E-ISSN*. pp. 2348-4470
- [13] Boehm, B., Abts, C., Brown, A., Chulani, S., Clark, B., Horowitz, E., Madachy, R., Reifer, D. & Steece, B. Software cost estimation with COCOMO II. (Prentice Hall Press,2009)
- [14] Manikavelan, D. & Ponnusamy, R. Software quality analysis based on cost and error using fuzzy combined COCOMO model. *Journal Of Ambient Intelligence And Humanized Computing*. pp. 1-11 (2020)
- [15] Nandal, D. & Sangwan, O. Software cost estimation by optimizing COCOMO model using hybrid BATGSA algorithm. *International Journal Of Intelligent Engineering And Systems*. **11**, 250-263 (2018)
- [16] Shukla, S. & Kumar, S. Applicability of neural network based models for software effort estimation. *2019 IEEE World Congress On Services (SERVICES)*. **2642** pp. 339-342 (2019)
- [17] Lavingia, K. & Mehta, R. Information retrieval and data analytics in internet of things: current perspective, applications and challenges. *Scalable Computing: Practice And Experience*. **23**, 23-34 (2022)
- [18] Sarro, F., Petrozziello, A. & Harman, M. Multi-objective software effort estimation. *2016 IEEE/ACM 38th International Conference On Software Engineering (ICSE)*. pp. 619-630 (2016)
- [19] BaniMustafa, A. Predicting Software Effort Estimation Using Machine Learning Techniques. *2018 8th International Conference On Computer Science And Information Technology (CSIT)*. pp. 249-256 (2018)
- [20] Anandhi, V. & Chezian, R. Regression techniques in software effort estimation using cocomo dataset. *2014 International Conference On Intelligent Computing Applications*. pp. 353-357 (2014)
- [21] Yeh, T. & Deng, S. Application of machine learning methods to cost estimation of product life cycle. *International Journal Of Computer Integrated Manufacturing*. **25**, 340-352 (2012)
- [22] Osareh, A. & Shadgar, B. Machine learning techniques to diagnose breast cancer. *2010 5th International Symposium On Health Informatics And Bioinformatics*. pp. 114-120 (2010)
- [23] Rahikkala, J., Hyrynsalmi, S., Leppänen, V. & Porres, I. The role of organisational phenomena in software cost estimation: A case study of supporting and hindering factors. *E-Informatica Software Engineering Journal*. **12**, 167-198 (2018)
- [24] Rong, S. & Bao-Wen, Z. The research of regression model in machine learning field. *MATEC Web Of Conferences*. **176** pp. 01033 (2018)
- [25] Drucker, H., Burges, C., Kaufman, L., Smola, A. & Vapnik, V. Support vector regression machines. *Advances In Neural Information Processing Systems*. **9** (1996)
- [26] Xu, S., An, X., Qiao, X., Zhu, L. & Li, L. Multi-output least-squares support vector regression machines. *Pattern Recognition Letters*. **34**, 1078-1084 (2013), <https://www.sciencedirect.com/science/article/pii/S0167865513000196>
- [27] Joy, R., Lavingia, A. & Lavingia, K. Performance evaluation of transmission distance and bit rates in inter-satellite optical wireless communication system. *Int J Adv Technol Eng Sci (IJATES)*. **4**, 36-40 (2021)
- [28] Zhong, H., Wang, J., Jia, H., Mu, Y. & Lv, S. Vector field-based support vector regression for building energy consumption prediction. *Applied Energy*. **242** pp. 403-414 (2019), <https://www.sciencedirect.com/science/article/pii/S0306261919304878>
- [29] Breiman, L. Random forests. *Machine Learning*. **45**, 5-32 (2001)
- [30] Schonlau, M. & Zou, R. The random forest algorithm for statistical learning. *The Stata Journal*. **20**, 3-29 (2020) (2019)

- [31] Ho, T. Random decision forests. *Proceedings Of 3rd International Conference On Document Analysis And Recognition*. **1** pp. 278-282 (1995)
- [32] Quinlan, J. & Others Learning with continuous classes. *5th Australian Joint Conference On Artificial Intelligence*. **92** pp. 343-348 (1992)
- [33] Al Asheeri, M. & Hammad, M. Machine learning models for software cost estimation. *2019 International Conference On Innovation And Intelligence For Informatics, Computing, And Technologies (3ICT)*. pp. 1-6

*Edited by:* Katarzyna Wasielewska-Michniewska

*Regular paper*

*Received:* Apr 1, 2023

*Accepted:* Jan 8, 2024

---

## AIMS AND SCOPE

The area of scalable computing has matured and reached a point where new issues and trends require a professional forum. SCPE will provide this avenue by publishing original refereed papers that address the present as well as the future of parallel and distributed computing. The journal will focus on algorithm development, implementation and execution on real-world parallel architectures, and application of parallel and distributed computing to the solution of real-life problems. Of particular interest are:

**Expressiveness:**

- high level languages,
- object oriented techniques,
- compiler technology for parallel computing,
- implementation techniques and their efficiency.

**System engineering:**

- programming environments,
- debugging tools,
- software libraries.

**Performance:**

- performance measurement: metrics, evaluation, visualization,
- performance improvement: resource allocation and scheduling, I/O, network throughput.

**Applications:**

- database,
- control systems,
- embedded systems,
- fault tolerance,
- industrial and business,
- real-time,
- scientific computing,
- visualization.

**Future:**

- limitations of current approaches,
- engineering trends and their consequences,
- novel parallel architectures.

Taking into account the extremely rapid pace of changes in the field SCPE is committed to fast turnaround of papers and a short publication time of accepted papers.

---

## INSTRUCTIONS FOR CONTRIBUTORS

Proposals of Special Issues should be submitted to the editor-in-chief.

The language of the journal is English. SCPE publishes three categories of papers: overview papers, research papers and short communications. Electronic submissions are preferred. Overview papers and short communications should be submitted to the editor-in-chief. Research papers should be submitted to the editor whose research interests match the subject of the paper most closely. The list of editors' research interests can be found at the journal WWW site (<http://www.scpe.org>). Each paper appropriate to the journal will be refereed by a minimum of two referees.

There is no a priori limit on the length of overview papers. Research papers should be limited to approximately 20 pages, while short communications should not exceed 5 pages. A 50–100 word abstract should be included.

Upon acceptance the authors will be asked to transfer copyright of the article to the publisher. The authors will be required to prepare the text in  $\text{\LaTeX} 2_{\epsilon}$  using the journal document class file (based on the SIAM's `siamltex.clo` document class, available at the journal WWW site). Figures must be prepared in encapsulated PostScript and appropriately incorporated into the text. The bibliography should be formatted using the SIAM convention. Detailed instructions for the Authors are available on the SCPE WWW site at <http://www.scpe.org>.

Contributions are accepted for review on the understanding that the same work has not been published and that it is not being considered for publication elsewhere. Technical reports can be submitted. Substantially revised versions of papers published in not easily accessible conference proceedings can also be submitted. The editor-in-chief should be notified at the time of submission and the author is responsible for obtaining the necessary copyright releases for all copyrighted material.