

# Scalable Computing: Practice and Experience

---

Scientific International Journal  
for Parallel and Distributed Computing

ISSN: 1895-1767



Volume 24(3)

September 2023

---

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 24, Number 3, September 2023

---

## TABLE OF CONTENTS

PAPER IN THE SPECIAL ISSUE ON INTELLIGENT CLOUD TECHNOLOGIES ENABLED SOLUTIONS FOR NEXT GENERATION SMART CITIES :

**Synchronous Federated Learning based Multi Unmanned Aerial Vehicles for Secure Applications** 191

*Itika Sharma, Sachin Kumar Gupta, Ashutosh Mishra, Shavan Askar*

PAPERS IN THE SPECIAL ISSUE ON CLOUD COMPUTING FOR INTELLIGENT TRAFFIC MANAGEMENT AND CONTROL:

**Optimization Algorithm for Urban Rail Transit Operation Scheduling based on Linear Programming** 203

*Shuang Wu, Jinlong Wu, Yifeng Sun, Tong Yao*

**Method for Identifying Motor Vehicle Traffic Violations Based on Improved YOLOV Network** 217

*Zhengjun Hao*

**Application of Improving ABC in Cold Chain Low Carbon Logistics Path Planning** 229

*Xiazu Bai*

**Time Window Oriented IoT Vehicle Pathway Study for the Dynamically Changing Needs of E-Commerce Customers** 241

*Xin Chen*

**Protecting Data and Privacy: Cloud-based Solutions for Intelligent Transportation Applications** 257

*Surjit Singha, Ranjit Singha*

PAPERS IN THE SPECIAL ISSUE ON SENTIMENT ANALYSIS AND AFFECTIVE COMPUTING IN MULTIMEDIA DATA ON SOCIAL NETWORK:

**Improving Bert Model Accuracy for Uni-modal Aspect-Based Sentiment Analysis Task** 277

*Amit Chauhan, Rajni Mohana*

**Design of Sentiment Analysis Framework of Digital Media Short Text Based on Multi-pattern Sentiment Lexicon** 287

*Shuqin Lin*

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

<b>Aspect-based Text Classification for Sentimental Analysis using Attention Mechanism with RU-BiLSTM</b>	<b>299</b>
<i>Sandeep Yeliseti, Nellore Geethanjali</i>	
<b>A Deep LSTM-RNN Classification Method for Covid-19 Twitter Review Based on Sentiment Analysis</b>	<b>315</b>
<i>Jatla Srikanth, Avula Damodaram Shanmugam</i>	
<b>Deep Learning-Based Education Decision Support System for Student E-learning Performance Prediction</b>	<b>327</b>
<i>Sudha Prathyusha Jakkaladiki, Martina Janečková, Jan Krunčák, Filip Malý and Tereza Otčenášková</i>	
<b>Flexible English learning platform using collaborative Cloud-Fog-Edge Networking</b>	<b>339</b>
<i>Sian Chen, Linqiang Tang</i>	
<b>Analysis, Prediction and Classification of Skin Cancer using Artificial Intelligence – A Brief Study and Review</b>	<b>355</b>
<i>Madhavi Latha Pandala, S. Periyamayagi</i>	
<b>General Layout Planning Model of Landscape Ceramic Sculpture Based on NSGA - II Algorithm</b>	<b>371</b>
<i>Henan Feng, Liqun Zheng, Shuang Qiao</i>	
<b>Evaluation of Curriculum Implementation Effectiveness of Higher Vocational Education Based on Collaborative Filtering Algorithm</b>	<b>379</b>
<i>Hao Wang, Huiyan Li</i>	
<b>A Method for Online Monitoring Data Release of Composite Submarine Cable Based on Horizontal Federated Learning</b>	<b>389</b>
<i>Xinli Lao, Jiajian Zhang, Chuanlian Gao, Huakun Deng, Yanlei Wei, Zhenzhong Liu</i>	
<b>Real-time Monitoring Method of Insulation Status of Photoelectric Composite Submarine Cable based on Thermoelectric Coupling</b>	<b>399</b>
<i>Xinli Lao, Jiajian Zhang, Chuanlian Gao, Huakun Deng, Yanlei Wei, Zhenzhong Liu</i>	
<b>The Optimal Selection of Ideological and Political Issues in Business Courses Based on Swarm Intelligence Algorithm</b>	<b>409</b>
<i>Xuecai Yin</i>	
<b>Intelligent Element Layout Algorithm of Ceramic Design Drawing Based on Genetic Algorithm</b>	<b>419</b>
<i>Xinqing Li, Juan Li, Chuanlian Gao</i>	

<b>Distribution Network Planning Method Considering the Coupling of Transportation Network and Distribution Network</b>	<b>429</b>
<i>Zhuoran Song, Jianfeng Li, Tao Jiang, Sichen Lu</i>	
<b>Remote Monitoring System of Digital Agricultural Greenhouse Based on Internet of Things</b>	<b>439</b>
<i>Lu Lu</i>	
<b>A Hybrid Education Quality Assessment Method based on Heuristic Optimization Algorithm</b>	<b>449</b>
<i>Lan Cui, Yuefen Zhi</i>	
<b>Using Machine Learning Algorithms to Design Personalized Exercise Programs for Health and Wellness</b>	<b>463</b>
<i>Yan Lu</i>	
<b>Analyzing Spectator Emotions and Behaviors at Live Sporting Events using Computer Vision and Sentiment Analysis Techniques</b>	<b>475</b>
<i>Yang Xu</i>	
<b>Scalability and Sustainability in the Construction of a Social Sports Management Information Platform based on Web Technology</b>	<b>487</b>
<i>Mengyan Wu, Yuanyuan Hou, Yueqin Tang</i>	
<b>Reinforcement Learning-based Algorithms for Music Improvisation and Arrangement in Sensor Networks for the Internet of Things</b>	<b>499</b>
<i>Xiaoling Hu</i>	
 PAPERS IN THE SPECIAL ISSUE ON NEXT GENERATION PERVASIVE RECONFIGURABLE COMPUTING FOR HIGH PERFORMANCE REAL TIME APPLICATIONS:	
<b>Vulnerability Detection in Computer Networks using Virtual Reality Technology</b>	<b>511</b>
<i>Songlin Liu</i>	
<b>Computer Malicious Code Signal Detection based on Big Data Technology</b>	<b>511</b>
<i>Xiaoteng Liu</i>	
 PAPERS IN THE SPECIAL ISSUE ON SCALABLE MACHINE LEARNING FOR HEALTH CARE: INNOVATIONS AND APPLICATIONS:	
<b>Big Data in Healthcare – A Comprehensive Bibliometric Analysis of Current Research Trends</b>	<b>531</b>
<i>Aijaz Ahmad Reshi, Arif Shah, Shabana Shafi, Majid Hussain Qadri</i>	

**Efficient Net-based Transfer Learning Technique for Facial Autism Detection** **551**

*Tariq Saeed Mian*

PAPER IN THE SPECIAL ISSUE ON MACHINE LEARNING AND BLOCK-CHAIN BASED SOLUTION FOR PRIVACY AND ACCESS CONTROL IN IOT:

**Privacy and Security Enhancement of Smart Cities using Hybrid Deep Learning-enabled Blockchain** **561**

*Joseph Awotunde, Tarek Gaber, L.V. Narasimha Prasad, Sakinat Oluwabukonla Folorunso, Vuyyuru Lakshmi Lalitha*

PAPERS IN THE SPECIAL ISSUE ON SCALABLE COMPUTING IN ONLINE AND BLENDED LEARNING ENVIRONMENTS: CHALLENGES AND SOLUTIONS:

**The Effect of Online and Offline Sports Safety Education combined with MOOC Platforms in Physical Education Teaching in Colleges and Universities** **585**

*Yuan Gao*

**Scalable Computing Infrastructure for Online and Blended Learning Environments** **597**

*Liao Xin*

**Application of Improved Apriori Algorithm in Innovation and Entrepreneurship Engineering Education Platform** **609**

*Xuanyuan Wu, Yi Xiao, Anhua Liu*

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

**Spacecraft Test Data Integration Management Technology based on Big Data Platform** **621**

*Nanqi Gong*

**Research on Space Image Fast Classification Based on Big Data** **631**

*Yunyan Wang, Peng Chen*

**Application of Data Visualization Interaction Technology in Aerospace Data Processing** **641**

*Jia Lin, Bigeng Zheng, Zhijian Chen*



## SYNCHRONOUS FEDERATED LEARNING BASED MULTI UNMANNED AERIAL VEHICLES FOR SECURE APPLICATIONS

ITIKA SHARMA\*, SACHIN KUMAR GUPTA†, ASHUTOSH MISHRA‡ AND SHAVAN ASKAR§

**Abstract.** Unmanned Aerial Vehicles (UAVs), also known as drones, have rapidly gained popularity due to their widely employed applications in various industries and fields, including search and rescue, agriculture, industry, military operations, safety, and more. Additionally, drones assist with tasks such as search and rescue efforts, pandemic virus containment, crisis management, and other critical operations. Due to their unique capabilities in image, video, and information collection, a multi-UAV system plays a crucial role in these activities. However, such images and video data involve individual privacy. Therefore, such multi-UAV applications have an indigenous tradeoff of privacy preservation. We have proposed a Federated Learning (FL) based approach for ensuring privacy in multi-UAV applications. The proposed methodology utilizes a synchronous FL approach and the Convolutional Neural Network (CNN) to ensure security. The model parameters are protected by using a secure aggregation. Results demonstrate that the proposed approach outperforms existing techniques in terms of accuracy and precision.

**Key words:** Federated Learning, Machine Learning, Privacy, Security, Unmanned Aerial Vehicles

**1. Introduction.** Unmanned Aerial Vehicles (UAVs) or drones, have experienced tremendous growth in both the industrial and academic fields. Drones play a crucial role due to their unique characteristics and ability to perform critical tasks such as capturing images and videos. The market value of UAV usage in civil infrastructures exceeds 45 billion dollars. A drone, also known as a UAV, is an aircraft that operates without an onboard operator or crew. Drones were first used in Italy in 1849 during the struggle for independence between Venice and Austria. Austrian soldiers attacked Venice using bomb-laden hot-air, hydrogen, or helium-filled balloons. In 1982, the Israeli Air Force utilized UAVs to destroy the Syrian fleet while minimizing Israeli losses [1]. Israeli UAVs were employed as a deception tactic, disrupting communications and providing real-time video surveillance. The rapid adoption of drones over the past decade has raised concerns about privacy, security, and safety. Drones are used by travelers and paparazzi to capture images of individuals in their homes and other previously private locations. Additionally, drones are utilized in high-risk areas such as cities and airports. UAVs are an integral part of an Unmanned Aircraft System (UAS), which includes a controller and communication system for the UAV. Companies like Google and Amazon are developing drones to facilitate the delivery of goods by air. Another intriguing concept being explored by Facebook involves the construction of massive drones to provide direct internet access to remote areas. Lightweight synthetic structures are commonly used in the construction of unmanned aircraft to enhance maneuverability and manage weight [2]. Military drones, thanks to the strength of composite materials, are capable of flying at extremely high altitudes. The combination of drones and IoT technology has resulted in a surge of new commercial applications. UAVs integrated with IoT sensor networks on the ground can assist agricultural organizations in monitoring lands and harvests, enable energy companies to inspect power lines and operate equipment, and help insurance companies assess claims and policies.

---

\*School of Electronics and Communication Engineering, Shri Mata Vaishno Devi University, Kakryal-182320, Katra, (Jammu and Kashmir), UT, India ([itikasharma142@gmail.com](mailto:itikasharma142@gmail.com))

†Department of Electronics and Communication Engineering, Central University of Jammu, Samba, UT of J&K, 181143 India and School of Electronics and Communication Engineering, Shri Mata Vaishno Devi University, Katra-182320, (UT of J&K), India ([sachin.gupta@smvdu.ac.in](mailto:sachin.gupta@smvdu.ac.in), Corresponding Author)

‡School of Integrated Technology, Yonsei, University, South Korea, and the Department of Electronics & Communication Engineering, Graphic Era Deemed to be University, Dehradun 248002, Uttarakhand, India ([ashutoshmishra@yonsei.ac.kr](mailto:ashutoshmishra@yonsei.ac.kr), Corresponding Author)

§Technical Engineering College, Information System Engineering Department, Erbil, Iraq ([shavan.askar@epu.edu.iq](mailto:shavan.askar@epu.edu.iq))

ML techniques like differential privacy and homomorphic encryption have surged in IoT devices. Even commonplace items used by people in daily life, such as wearables, smart meters, and smart water meters, can be utilized with the Internet of Things (IoT) network to collect data from the environment using sensors and actuators. This data can then be used to solve a variety of problems in daily life. IoT has transformed traditional industries into smart ones, including smart waste management, smart grids, smart homes, and smart healthcare [3]. However, the heterogeneity of devices, energy shortages, restricted bandwidth, and other limitations pose challenges in implementing advanced security methods in IoT networks, making them more susceptible to destructive insider attacks [3]. Also, security problems lead to privacy-sensitive issues. There are security issues in IoT networks because central data is collected to train the model. So, protecting data is quite critical as there is a central server in a common model that works on training and test sets.

There are many challenges in UAV civil applications, including swarming challenges and network security challenges. Another issue is the scalability issue [4]. Therefore, the unacceptable latency and the raw data transfer to the server by UAV requiring high network bandwidth are the most significant issues faced by traditional ML. To handle these problems, we suggest using federated learning (FL) model that efficiently addresses all privacy and latency issues. Google introduced the first application of FL in 2016. They established FL to improve the predictive text on Google's Android keyboard. Apple also utilizes FL to enhance voice recognition on Siri. FL can protect clients' raw data and reduce communication costs. There is widespread use of FL in the healthcare system, where it securely handles data without sending raw data directly to the server [5, 6, 7, 8]. FL utilizes less complex hardware as it doesn't rely on a central server. Each UAV locally trains its models, and then the updated parameters are sent to the server and aggregated globally there [9, 10, 11, 12, 13, 14].

**1.1. Key Contributions.** The FL methodology has been used to address multi-UAV systems due to security concerns associated with DL-based methodologies that transmit raw data. FL was developed as a solution to this problem, aiming to store original data at its source and solely send locally trained models from users to the server for aggregation. Both asynchronous and synchronous UAVs are employed in conjunction with federated learning. Asynchronous federated learning allows for faster training of local models, but it entails a loss of data packets, resulting in subpar communication and diminished accuracy. Therefore, we utilize federated learning to effectively communicate and achieve accuracy. The key contributions of this study are as follows:

1. Application of FL in conjunction with multi-UAV systems to ensure privacy.
2. Utilization of convolutional neural networks (CNN)-based algorithm for model training.
3. Comparative analysis of the traditional centralized model versus the federated model.
4. Performance assessment of the federated models in terms of accuracy, precision, and recall.

The remainder of the paper is structured as follows: part II reviewed related literature. The proposed model is illustrated in Section III. The debate and findings are examined in Section IV. Section V serves as the work's conclusion with its future scope.

**2. Related Works.** This section will differentiate the study of previous work done from our article by revealing the procurement of UAVs with FL. Al-Emadi et al. in [1] discussed a drone detection solution based on RF frequency. The author proposed a model for detecting and classifying drones using an RF signal and compared CNN and DNN. In this model, the author used a DL technique called CNN, which is an effective mechanism for drone detection. The proposed model was tested and trained using a dataset. To analyze and determine the performance of this model, the author made a comparison of data obtained with ANN, CFAR, and HOC. Due to its unique qualities that integrate feature extraction and classification into a single model, CNN has a wide range of applications, particularly in recognition and classification challenges. In this paper, the author claims to have achieved accuracy and provided better performance in drone detection as well as drone identification.

Brik et al. in [2] introduced the FL concept and its fundamentals. The author gives a broad overview of FL applications for UAV networks. The traditional DL methods use a centralized server to which the UAV data is directly sent. This creates issues in network communication overhead, leading to energy inefficiency and network bandwidth problems for UAV devices. Moreover, there may be private data like localization and identification of UAVs, which can directly impact UAV privacy. To address these concerns, the author



introduced the FL concept, where data is not directly sent to the server. The basic concept behind UAVs is to preserve raw data where it is created and locally train the models. In this paper, the author discusses the open issues and challenges in the FL approach.

In [3], Ahmad et al. demonstrated malicious attacks in the IoT using supervised machine learning techniques. The IoT was created to connect numerous smart gadgets and sensors for gathering and analyzing vital information online. However, the IoT network faces several limitations, including limited sensor computational capability, heterogeneity of gadgets, limited energy resources, and bandwidth, among others. These limitations prevent the use of high-end security techniques, making the networks more susceptible to insider assaults and other security threats. Additionally, due to the unexpected behavior and ubiquity of IoT networks, identifying malevolent insiders in the network is exceedingly challenging. Machine learning techniques can address these issues by forecasting anomalies in the system and understanding its behavior. The authors applied various supervised machine-learning approaches to the available IoT dataset to determine the most effective method for identifying hostile insider assaults in the IoT network. Multiple supervised algorithms were trained on the NSL-KDD dataset, and the results show that the Gradient Boosting technique outperforms the other covered techniques.

The FL-based approach is the suggestion made by Lim et al. [4] to enable security in IoV applications such as traffic forecasting. In this article, the authors have developed an FL-based detecting and collaborative learning scheme in which UAVs gather data and participate in private information collaborative model training for IoV paradigm applications, aiming to establish an Intelligent Transportation System (ITS). Additionally, the authors presented a multidimensional contrast matching-based reasoning framework layout that aims to utilize the self-revealing characteristics of an ideal contract to produce the most ideal UA. Yang et al. [5] suggested an asynchronous FL framework for multi-drone networks, which may provide asynchronous learning or distributed computing by allowing the training model to operate without transmitting or sending raw sensitive information to UAV servers. It also includes a device selection technique to prevent devices of poor quality from compromising efficiency and precision in learning. Furthermore, the author has proposed an asynchronous advantage actor-critic (A3C)-based combined device selection, UAV deployment, and resource management strategy to increase federated convergence speed and accuracy. Simulation results reveal that the suggested scheme and algorithm produce greater accuracy and quicker federation execution time than existing solutions.

Yang et al. demonstrated energy-efficient UAVs based on federated learning [6]. This article examines the problem of resource allocation for federated learning across UAVs and efficient energy transmission. With the goal of reducing the overall energy consumption of the devices while considering the delay restriction, both the challenges of local processing and transmission energy are approached as optimization problems. The author presents a solution to the energy minimization problem by providing the delivery time for the reduction problem and suggests a bisection-based strategy to obtain the optimal answer, which is a practical option due to the iterative technique recommended by the author. Khamidehi et al. proposed a federated-based architecture in order to shorten UAV flight times while maintaining reliable internet connectivity [7]. Since UAVs are unable to independently develop the global model, the author provides the following two options: (1) Using federated learning, the UAVs cooperate to build a large model to predict outages in the environment; (2) Based on a model created in the first phase using fast random trees (RRTs), the author provides a path-planning technique. The author devises an optimal course for UAVs that reduces travel time while ensuring cellular connectivity. In [8], Yazdinejad et al. proposed a federated learning approach based on deep neural networks (DNN) for drone authentication. By integrating drones with federated learning, maximum privacy is achieved. The author creates an authentication model based on the DNN Neural Network and the Federated Learning idea. This authentication mechanism federates the number of aircraft while maintaining anonymity. This approach protects privacy and utilizes DNN design to scale with a large number of drones. Secure aggregation techniques and cryptographic algorithms are used to safeguard secure parameters for the models. The author states that the simulation demonstrates a high True Positive rate and performs best when used with the training and validation datasets.

Tang et al. conducted research [9] to improve Federated Edge Learning (FEEL) for B5G/6G networking in the Internet of Things (IoT) enabled by UAVs. Federated learning is an effective framework for creating a decentralized shared paradigm between computers and edge devices without transmitting raw data. However,

FEEL's performance in connected devices with UAV support is limited by two important factors: latency and energy consumption. While earlier research has focused on reducing latency and promoting energy efficiency, few studies have examined how the limited battery life of devices affects FEEL.

Wang et al. proposed an asynchronously federated learning system based on a permissioned blockchain in [10]. The article suggests a blockchain-based synchronization federated learning server that consists of a core blockchain and several sub-blockchains. Each sub-blockchain is responsible for updating a specific model parameter while the main blockchain updates all parameters globally. Based on this design, a permissioned blockchain technology federated learning synchronization aggregation protocol is suggested. This protocol combines second-order aggregation calculations with the learning algorithm to effectively replace concurrent federated supervised learning. As a result, synchronization problems are avoided and the reliability of distributed data is guaranteed. According to calculations and experimental results, the suggested architecture can maintain acceptable training results when working with a small number of nodes with varying data quality, has good automated failover, and can be expanded to edge computing applications.

In [11, 12, 13], the authors proposed a privacy computing system (AFLPC) for 5G-V2X scenarios using asynchronous federated learning to address potential privacy leakage issues at the network edge in the 5G V2X environment. This model combines secure computing and asynchronous federated learning. The article introduces an adaptive independent privacy strategy to protect parameter privacy and reduce the impact of noise on model validity. To address the consolidation issue of asynchronously federated learning based on moving averages in practical applications, an aggregate approach based on weighted summaries is recommended. The proposed asynchronous federated instruction and learning mechanism primarily focuses on the 5G V2X environment, but further investigation into other settings is also suggested [14, 15, 16, 17, 18]. Results from the study conducted using the MNIST and CIFAR-100 datasets demonstrate that the approach described in this article can provide high precision in terms of confidentiality and safety, solve the issue of diminishing model calculation time caused by variations in node learning speed, and improve the productivity of AFL.

The use of "UAV with FL" demonstrates that when UAVs are used with federated learning, it ensures the privacy of transmitted data without any loss and avoids communication overhead. This leads to efficient network bandwidth and energy utilization. In traditional DL-based UAVs, the original information is sent to a central server, which creates issues related to network bandwidth, energy efficiency, and privacy. Due to these concerns, the concept of federated learning is now being used, where only local models are transferred instead of the original information. This approach offers the benefits of privacy, energy efficiency, and network bandwidth.

**3. FL: A System Model.** This section covers the proposed model, the function of FL in UAVs, the function of CNN, and the encryption technique utilized in the FL process.

**3.1. Proposed Model.** The majority of machine learning (ML) techniques are cloud-based, which means that data must be transferred to and analyzed in a single location, such as a database server. However, these ML methods are not suitable for UAV-wireless networks due to several reasons. Firstly, since the generated data could contain sensitive information like the location and identification of UAVs, the private data might not be accessible. Secondly, the continuous transmission of unprocessed data by UAVs, such as images, videos, and other types of data, to the server requires a high network bandwidth and consumes more UAV energy, especially when bandwidth and UAV energy are constrained.

Moreover, cloud-centric approaches introduce significant latency, particularly for applications that require instantaneous decisions, like unmanned drone monitoring and UAV-based virtual reality applications. Therefore, decentralized teaching methods are crucial for effectively managing scattered sub-datasets produced by UAV instruments. In 2016, Google introduced Federated Learning (FL) as a novel form of AI that moves training to the edge or on-device. FL is based on decentralized data and training, enabling multiple parties to collaborate on a shared, robust ML model without disclosing data. It addresses critical issues related to data confidentiality, safety, data login credentials, and access to heterogeneous data.

In our proposed model, we incorporate privacy by deploying multiple UAVs and utilizing FL. The server sends the model to each UAV for training, and these neighboring UAV clients regularly exchange the model parameters. Before delivering these parameters, several models encrypt them, thus enhancing data security and protection. The features of the global model are sent to data centers to include them in their ML model

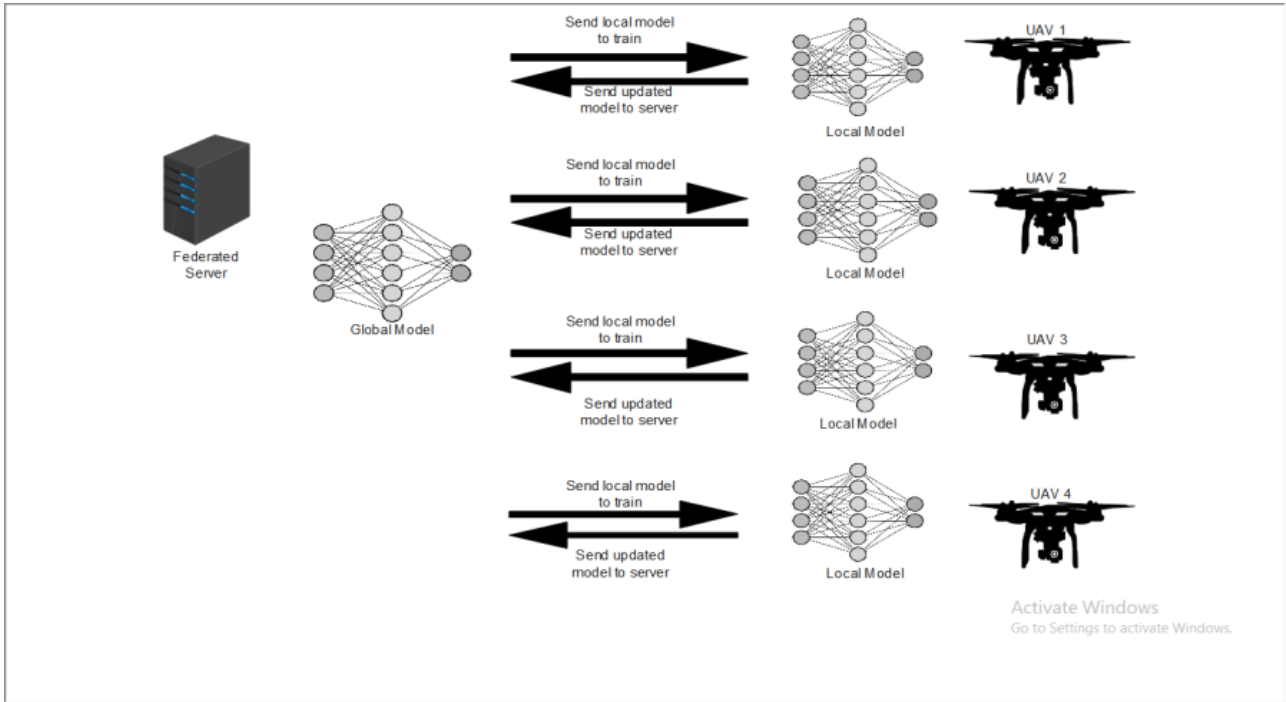


Fig. 3.1: FL Process

parameters. Each UAV then trains the model on its own data before sending the revised model back to the server. The models are globally aggregated at the server, and the process is repeated until the desired accuracy is achieved. As a result, privacy is ensured as raw data is not sent directly to the server; only models are transmitted. Secure aggregation, which encrypts both the data and the model, is employed in our work. Secure aggregation is essentially a protocol used for privacy-preserving federated learning. It enables the aggregation of models without revealing individual models, thereby preventing the disclosure of private information from both individual and local models. This aggregation ensures that the server cannot access the private data of the local models. Additionally, Fig. 3.1 illustrates the federated learning process. First, the FL server creates the global model  $G_0$  and sends the UAVs the requirements for data types and training parameters. To prevent the exhaustion of the UAV's resources, the FL server displays the number of iterations and the learning rate. For each UAV  $i$  start gathering information and updating local parameters,  $L_{ij}$ . Using the global model  $G_j$  ( $j$  is the iteration index), UAVs also search for the optimal settings to minimize the loss function. The modified settings are then provided to the server. The FL server compiles the local updates and sends the refreshed parameters to the UAVs. The purpose of the FL server is to minimize the overall global loss function ( $G_j$ ) as shown in equation (3.1):

$$Loss(G^j) = \frac{1}{M} \sum_{i=1}^{=M} Loss(L_i^j) \quad (3.1)$$

In the initial stage of implementation, each UAV receives models from the server. The initial stage of the suggested methodology's implementation includes sending models to every local UAV client. The second phase of implementation will then begin after the training procedure has been finished. Each UAV client begins the training process using their own data in the second phase. The trained model will then be sent back to the drones in a third stage when there is no or low accuracy from the server and federated server. The procedure will continue until the correctness of the data is maintained. After the servers have globally combined the federated learning model and accuracy is maintained, the federated server will broadcast the globally aggregated model.

**3.2. Role of FL in UAVs.** Security risks are increasing every day as the number of UAVs rises. Each day, a large number of drones are being attacked. Therefore, there is a critical demand for highly secure drones. Several researchers have developed AI-based drones and incorporated privacy into them using the concepts of ML and DL. However, there have still been some issues with achieving truly secure communication. To address this problem, the concept of FL was introduced.

FL is an ML method that involves training an algorithm on numerous decentralized edge devices or servers. Each device or server retains local data samples and does not share them. In contrast, conventional DL-assisted systems are centralized and rely on servers to receive and store UAV data. However, UAVs are resource-constrained devices, particularly in terms of CPU and electrical resources. By avoiding the transfer of data to a central organization, FL significantly reduces network overhead. Consequently, FL utilizes less bandwidth compared to centralized ML systems.

This implies that FL may enable the training of learning models for UAV wireless communication in the future, in contrast to cloud-centric approaches that are centralized. FL avoids transmitting sensitive information to a single node, ensuring the security of UAV data while also reducing network costs and latency. In our work, we focused on the CNN model for the training phase. Although any NN can be used for training, the CNN model offers advantages in terms of computational efficiency, memory efficiency, and complexity efficiency. FL is increasingly being adopted in various domains, including healthcare, autonomous vehicles, mobile applications, predictive maintenance in industries, and others.

**3.3. Role of CNN Model.** A Convolutional Neural Network (CNN) is constructed with a convolutional layer that performs a process known as convolution. During the convolution process, the input undergoes convolution with filters, leading to activation. CNN has a broad range of applications in object recognition and is primarily employed in image classification to evaluate visual information. CNN can be utilized for tasks such as image and video categorization, identification, medical image analysis, and recommendation systems. Convolutional, ReLU, Pooling, and Fully Connected (FC) layers are the concealed layers within a CNN.

**3.4. Homomorphic Encryption.** Rivest et al. began their exploration of the creation of the homomorphic encryption (HE) method in 1978. HE is a unique encryption method that addresses security and privacy concerns. One option is to encrypt both our models and data. This encryption technique allows data to remain encrypted while being processed and handled. Model security is ensured during model training. HE is an encryption technology that enables computations to be performed on encrypted information without the need for a secret key to decrypt it.

The HE approach can be used to construct privacy-preserving machine learning (PPML) in industries where privacy protection is important. HE is an encryption technology that empowers service providers to perform calculations. Network operators can directly execute arithmetic computations on ciphertexts by utilizing HE encryption technology. Partially homomorphic encryption (P.H.E.) was originally capable of addition and multiplication.

**3.5. Practical Applications of Federated Learning in UAV.** The following are potential real-time applications of FL-enabled UAVs:

- A2G photographs can sometimes be taken and kept private, which is why FL-based UAVs come into service. FL-based UAVs keep data safer because the initial information is not sent directly to the server. As a result, no one will be able to easily breach the data area.
- FL-based UAVs play a significant role in data security and privacy, especially in the military. Given the sensitivity of military information, FL-based UAVs are essential for maintaining security and privacy.
- Federated technologies can be utilized to generate user behavioral patterns from a pool of smartphone data, including voice recognition, face identification, and next-word prediction, without disclosing any personal information. For example, Google employs federated learning to develop pattern recognition algorithms for on-device voice commands like "Hey Google" in Google Assistant.
- Manufacturing companies can utilize federated learning algorithms to create equipment predictive maintenance predictions. Predictive maintenance may face challenges when clients or users are reluctant to provide personal information, and when there are data exporting issues from various locations. Federated learning can overcome these problems by utilizing regional datasets.

Table 4.1: Simulation Parameters

Description	Value
UAV Servers	1
Epochs	60
UAV Clients	4
CPU Usage	2.5Ghz
Global Training Accuracy	97
Global Training Loss	3
RAM	16GB

- Federated learning techniques would be beneficial for insurance and healthcare systems as they allow for the protection of private data in its original form. Federated learning systems can gather data from multiple sources such as hospitals and electronic health record databases to diagnose uncommon diseases, thereby providing a more diverse set of data. In contemporary healthcare systems, hospitals, research institutions, and federal departments collaborate to enhance healthcare nationwide.

**3.6. Core Challenges of the Federated Learning Process.** A number of major obstacles must be overcome for Federated Learning to be implemented in UAVs:

- Effective communication between federated learning-enabled UAV networks.
- Managing diverse systems within a single network.
- Data in federated networks are statistically heterogeneous.
- Privacy issues and methods to protect them.

**4. Results and Discussions.** In this section, the gathered results discuss the training accuracy and training loss of local UAV client models, as well as the global model. The performance has been analyzed using Python libraries, including NumPy, TensorFlow, pandas, and Keras. Socket programming is used to communicate local client models to the server, and an HTTPS connection is established for the interconnection of multiple UAVs and servers. In this work, 6000 training samples are taken, and each sample is tested. The experimental parameters used throughout the work are described in Table 4.1.

**4.1. Flowchart of the Proposed Work.** Fig. 4.1 represents the flowchart of the proposed work. Firstly, the UAV client is loaded with the models or architecture, weights, and gradients. Then, the models undergo training using CNN. The input layer of the CNN takes the models or architectures as input and various hidden layers perform the learning process, resulting in updated or trained local models. The trained local models are encrypted using secure aggregation and homomorphic encryption and then sent to the server for global aggregation. Subsequently, the global models are broadcasted.

In Fig. 4.2 the training process of local UAV clients is depicted in Python, showing the training accuracy of the local models. During the training process, local model 4 shows accuracy after 10 epochs, while local models 1, 2, and 3 show accuracy after 40-45 epochs. The accuracy and loss graphs depend on the batch size and the type of data used. Some data is smooth and can be filtered easily, requiring less time for training. However, some data is more challenging to filter, so when a batch size of 200 is used, the graph reflects the corresponding accuracy and loss.

In Fig. 4.3 the training accuracy of the global model is displayed. This figure was generated using Python. The global model demonstrates accuracy after 10 epochs during the training process, while all other local clients (Aircraft) achieve stability after 30 iterations. Once the local clients are aggregated through the Federated Learning (FL) process, their accuracy and stability increase, while the remaining local clients exhibit different characteristics. Therefore, Figure 4 illustrates that local UAV clients achieve accuracy after 30 epochs, indicating that FL can achieve privacy in multiple ways.

Fig. 4.4 presents a comparison graph of accuracy between centralized and FL approaches. We utilized FL to train UAV data and used centralized learning with Python to train the same data, comparing the results. In the case of centralized and automated systems, Data 1, 2, and 4 indicate an accuracy of 80%, Data 5 has an

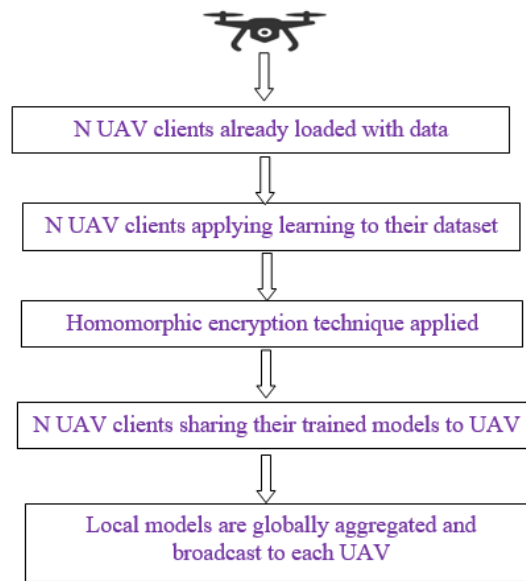


Fig. 4.1: Flowchart of Proposed Work

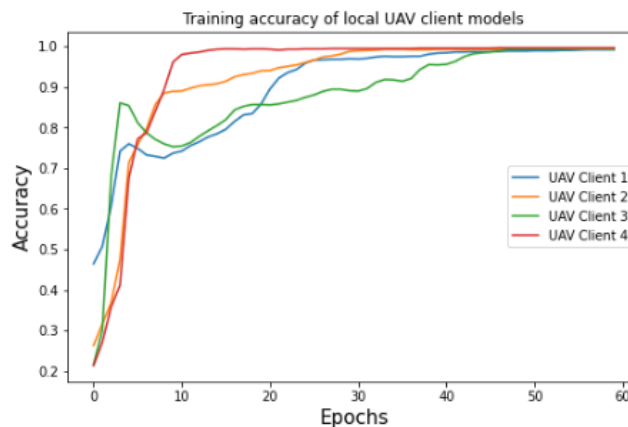


Fig. 4.2: Training Accuracy of Local Model

accuracy of 75%, and Data 3 has an accuracy of 90%. For FL, Data 1 and 4 demonstrate an accuracy of 82%, while Data 2, 5, and 3 provide accuracies of 80%, 75%, and 88%, respectively. The accuracy in FL is lower than in centralized learning, as evidenced by Data 3, suggesting that accuracy can depend on the type of data collected. Perhaps the dataset used in that case is not as clear and filtered.

Similarly, in Fig. 4.5, we conducted a comparison between centralized and FL approaches and depicted the graph. Data 1, 4, and 5 exhibit a precision of 80%, while Data 2 shows a precision of 85%, and Data 3 shows 90% in the case of centralized learning. For FL, Data 1 and 5 show a precision of 85%, Data 2 and 4 show a precision of 90%, and Data 3 shows a precision of 75%. In this graph, Data 1 and Data 5 show almost no variation in terms of precision. Data 3 displays a significantly lower precision rate in the case of FL compared to centralized learning. Thus, in terms of precision, FL yields result similar to centralized learning, with only minor variations observed in the comparative analysis.

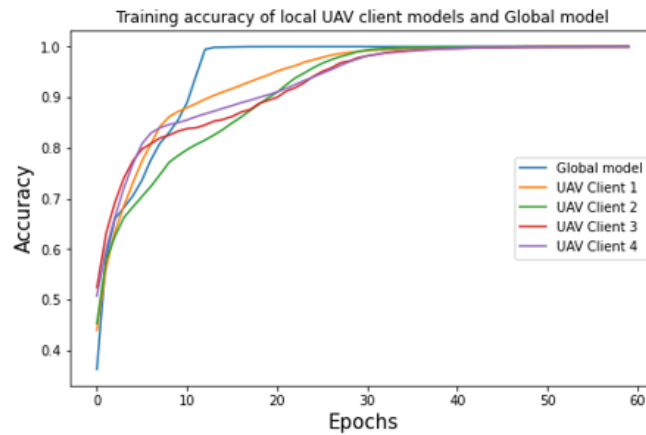


Fig. 4.3: Training Accuracy of the Global Model

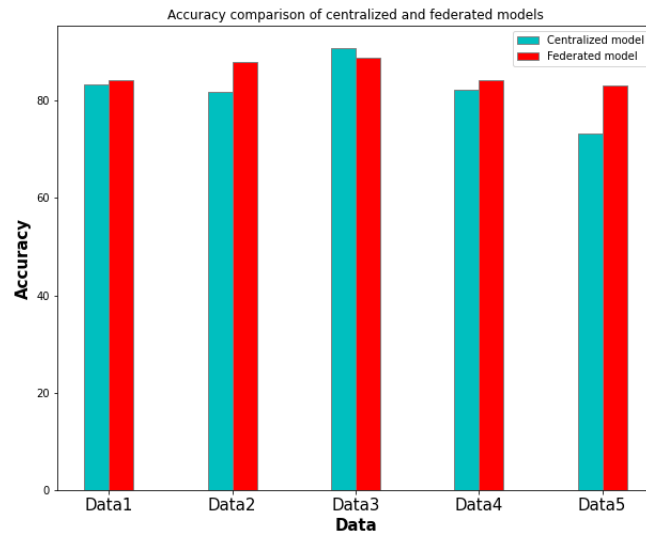


Fig. 4.4: Comparison Graph of the Accuracy of Centralized and Federated Learning

In Fig. 4.6, we employed UAV data and trained it using the FL process and then the centralized process, subsequently conducting a comparative analysis based on recall. Data 1 and 4 indicate the recall rate, Data 2 shows 80%, Data 3 shows a recall of 90%, and Data 5 shows 75% in the case of centralized learning. For FL, Data 1 and 2 demonstrate a recall of 90%, Data 3 shows 85%, and Data 4 and 5 exhibit a recall of 83%. In this case, as well, Data 3 shows better results in centralized learning and lower results in the case of FL. Data 4 displays significant variation in terms of recall for FL and centralized learning.

**5. Conclusion and Future Work.** The primary focus of this paper is the deployment of multi-UAVs and the incorporation of privacy measures through the utilization of FL. This study addresses machine learning-specific challenges, such as centralized training. The proposed model applies FL techniques to train each UAV's data locally and leverages decentralized UAV data to achieve privacy-preserving advantages. Furthermore,

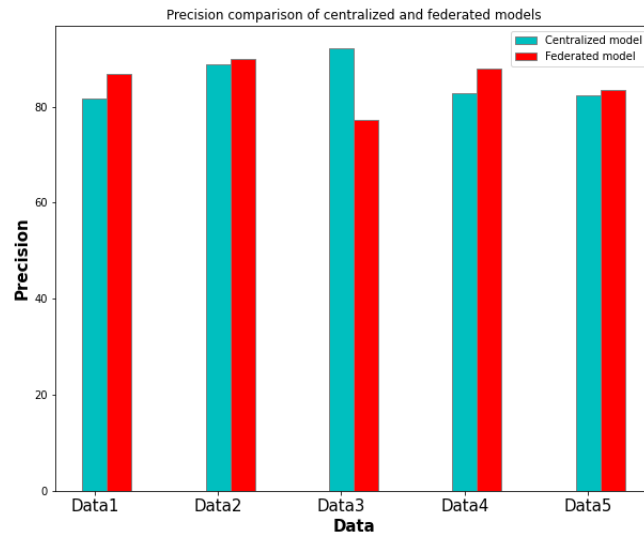


Fig. 4.5: Comparison Graph of Precision of Centralized and Federated Learning

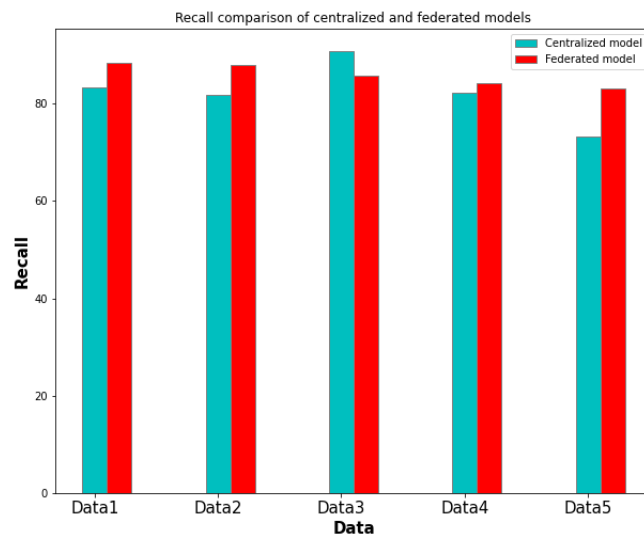


Fig. 4.6: Comparison Graph of Recall of Centralized and Federated Learning

this paper includes a comparative analysis between ML Models and FL Models, evaluating their accuracy, precision, and recall. The findings demonstrate FL's superiority over centralized learning regarding privacy protection, as raw data is not directly transmitted to the server. Instead, UAVs locally train their models, which are subsequently sent to the server for aggregation as a "global model." The results reveal enhanced accuracy, precision, and recall in the federated model compared to the local model, providing a comprehensive



comparison of traditional centralized versus federated models. Future research endeavors could expand upon the development of UAV networks to facilitate intercommunication among them. Additionally, the optimization of UAV placement in terms of height and angle could be explored to improve coverage, capacity, and secure communication within the network. Furthermore, the integration of FL into a multi-UAV network can enable the implementation of various efficient routing protocols and path planning techniques to enhance quality of service (QoS) metrics.

## REFERENCES

- [1] S. AL-EMADI, A. AL-ALI, A. MOHAMMAD, AND A. AL-ALI, *Audio Based Drone Detection and Identification using Deep Learning*, *15th International Wireless Communications and Mobile Computing Conference, 2019*, 459-464, doi:10.1109/IWCMC.2019.8766732.
- [2] B. BRIK, A. KSENTINI AND M. BOUAZIZ, *Federated Learning for UAVs-Enabled Wireless Networks: Use Cases, Challenges, and Open Problems*, *IEEE Access*, 8, 53841-53849, 2020, doi: 10.1109/ACCESS.2020.2981430.
- [3] M. S. AHMAD, AND S. M. SHAH, *Mitigating Malicious Insider Attacks in the Internet of Things using Supervised Machine Learning Techniques*, *Scalable Computing: Practice and Experience*, 22(1), 13–28, 2021, <https://doi.org/10.12694/scpe.v22i1.1818>.
- [4] W. Y. B. LIM, J. HUANG, Z. XIONG, J. KANG, D. NIYATO, X.-S. HUA, C. LEUNG, C. LEUNG, AND C. MIAO, *Towards federated learning in UAV- enabled Internet of Vehicles: A multi-dimensional contact-matching approach*, *CoRR abs/2004.03877*, 2020.
- [5] H. YANG, J. ZHAO, Z. XIONG, K. -Y. LAM, S. SUN, AND L. XIAO, *Privacy-Preserving Federated Learning for UAV-Enabled Networks: Learning-Based Joint Scheduling and Resource Management*, *IEEE Journal on Selected Areas in Communications*, 39(10), 3144-3159, Oct. 2021, doi:10.1109/JSAC.2021.3088655.
- [6] Z. YANG, M. CHEN, W. SAAD, C. S. HONG, AND M. S. BAHAI, *Energy Efficient Federated Learning Over Wireless Communication Networks*, *IEEE Access*, Nov 2020, (99):1-1 doi: 10.1109/TWC.2020.3037554.
- [7] B. KHAMIDEHI, AND E. S. SOUSA, *Federated Learning for Cellular-Connected UAVs: Radio Mapping and Path Planning*, *IEEE Global Communications Conference*, 2020, pp. 1-6, doi: 10.1109/GLOBECOM42002.2020.9322349.
- [8] A. YAZDINEJAD, R. M. PARIZI, A. DEGHANTANHA, AND H. KARIMPOUR, *Federated Learning for Drone Authentication*, *Adhoc Networks*, *Elsevier*, 120, 1-6, 102574, 2021, doi:10.1016/j.adhoc.2021.102574.
- [9] S. TANG, W. ZHOU, L. CHEN, L. LAI, J. XIA, AND L. FAN, *Battery constrained federated edge learning in UAV- enabled IoT for B5G/6G networks*, *Physical Communication*, *Elsevier*, 47, 1-6, August 2021, <https://doi.org/10.1016/j.phycom.2021.101381>.
- [10] R. WANG, AND W. T. TSAI , *Asynchronous Federated Learning System Based on Permissioned Blockchains*, *Sensors*, *MDPI*, 22(4), 1-18, 1672, 2022. <https://doi.org/10.3390/s22041672>.
- [11] J. HUANG, C. XU, Z. JI, S. XIAO, T. LIU, N. MA, AND Q. ZHOU, *An asynchronous, Federated learning privacy-preserving computing model applied to 5G-V2X*, *Security and Communication Networks*, *Hindawi*, 2022, 1-11, March 2022, doi:<https://doi.org/10.1155/2022/9334943>.
- [12] K. NEERAJA, AND G. NARSIMHA, *A Multi-Objective Hybrid Collision-free Optimal Path Finder for Autonomous Robots in Known Static Environments*, *Scalable Computing: Practice and Experience*, 23(4), 389–402, 2022, <https://doi.org/10.12694/scpe.v23i4.2049>.
- [13] C. ZHU, X. ZHU, J. REN, AND T. QIN, *Blockchain Enabled Federated Learning for UAV Edge Computing Network: Issues and Solutions*, *IEEE Access*, 10, 56591 - 56610, 13 May, 10.1109/ACCESS.2022.3174865.
- [14] B. TAHA, AND A. SHOUFAN, *Machine Learning-Based Drone Detection and Classification: State-of-the-Art in Research*, *IEEE Access*, 7, 138669-138682, 2019, doi:10.1109/ACCESS.2019.2942944.
- [15] A. BOGUETTAYA, AND H. ZARZOUR, A. KECHIDA, *Deep learning techniques to classify agricultural crops through UAV imagery: a review*, *Neural Computer and Application*, *Springer*, 34, 9511–9536, 2022.
- [16] M. SHAHEEN, M. S. FAROOQ, T. UMER, AND B.-S. KIM, *Applications of Federated Learning; Taxonomy, Challenges, and Research Trends*, *Electronics*, *MDPI*, 11(4), 1-6, February 2022, <https://doi.org/10.3390/electronics11040670>.
- [17] A. MISHRA, S. LEE, D. KIM, AND S. KIM, *In-Cabin Monitoring System for Autonomous Vehicles*, *Sensors*, *MDPI*, 1-21, 2022, 22, 4360. <https://doi.org/10.3390/s22124360>.
- [18] A. MISHRA, J. CHA, AND S. KIM, *Privacy-Preserved In-Cabin Monitoring System for Autonomous Vehicles*, *Deep Learning for Intelligent Surveillance Systems*, *Computational Intelligence and Neuroscience*, *Hindawi*, 1-15, 2022, 5389359, <https://doi.org/10.1155/2022/5389359>.

*Edited by:* Pradeep Kumar Singh

*Special issue on:* Intelligent Cloud Technologies Enabled Solutions for Next Generation Smart Cities

*Received:* Mar 7, 2023

*Accepted:* July 28, 2023





## OPTIMIZATION ALGORITHM FOR URBAN RAIL TRANSIT OPERATION SCHEDULING BASED ON LINEAR PROGRAMMING

Shuang Wu\* AND Jinlong Wu† AND Yifeng Sun‡ AND Tong Yao§

**Abstract.** At present, the traditional urban public transportation system cannot meet people’s daily travel needs. Urban Rail Transit (URT) has been rapidly promoted in major cities due to its advantages such as low energy consumption, high frequency, and large traffic volume. To achieve a more excellent and energy-saving operation scheduling strategy, the research first combines the train dynamics model and the energy consumption model. Since the optimization problem of URT is a linear problem, the attraction model of the Firefly algorithm can determine the calculation time consumed by the algorithm, which is very suitable for the complex optimization problem of URT. Therefore, the FA based optimization algorithm for urban rail transit operation scheduling (FURTOSO) based on the Firefly algorithm is studied and designed. Therefore, based on the study of the four working conditions of traction, cruise, coasting, and braking, a Firefly Algorithm for Urban Rail Transit Operation Scheduling (FURTOSO) was designed. Finally, the study optimizes the operation scheduling of Chengdu Metro Line 8 from two aspects: driving strategy and train schedule. The research demonstrates that the FURTOSO algorithm only needs 76 iterations to reach a stable state, with a fitness value of 0.6827. In practical applications, the utilization rate of train RBE is 30.1%, the total energy consumption (TEC) is  $2.661 * 10^{11}J$ , and the energy saving rate is 13.03%. In summary, the FURTOSO algorithm proposed in the study has excellent performance and has better energy-saving effects in Chengdu Metro Line 8.

**Key words:** Urban Rail Transit; Scheduling Optimization; Regenerative Braking; Firefly Algorithm; Energy Conservation

**1. Introduction.** As the boost of China’s urbanization, it is difficult for conventional public transportation systems to meet the rapidly growing passenger demand of cities [1]. At this time, urban rail transit (URT) emerged as the times require. It has advantages such as fast speed, large traffic volume, and environmental protection, which greatly alleviates urban congestion [2]. Currently, 40 cities in China have opened URTs, with a total of 212 lines constructed and a total operating mileage of 6730.27 kilometers. However, it has generated huge electricity consumption, with an estimated annual electricity consumption of 40 billion kilowatt hours, accounting for over 5% of the total national electricity consumption in the future. Therefore, studying the optimization of URT operation scheduling is of great significance. The power consumption of URT system mainly includes traction, ventilation, air conditioning, and lighting; Traction and ventilation and air conditioning account for 3/4 of the TEC, but there are still difficulties in studying energy conservation and emission reduction in some areas, such as ventilation and air conditioning systems [3]. It relies on developing efficient equipment and reducing working hours [4]. The energy consumption of traction power supply is closely related to the operating time under traction conditions. Therefore, the energy loss during train operation (TO) can be reduced through reasonable arrangement of driving strategies and schedules for TO. However, at present, there are few studies on the optimization of URT operation scheduling. The optimization of URT is a linear problem, and the Firefly algorithm (FA) can be used to solve the trust optimization problem with discrete variables after optimization by many scholars due to its simplicity and efficiency. In addition, as a new type of fully intelligent algorithm, the FA algorithm is mostly studied to solve simple optimization problems, and the use of this algorithm can make it adaptable to more fields. To solve the energy consumption problem of URT, the study starts with the utilization of regenerative braking energy and applies its FA algorithm to the scheduling optimization of multi train operation. Based on this, a FA based optimization algorithm for urban rail transit operation scheduling (FURTOSO) is constructed.

---

\*Department of Transportation Engineering, Yangzhou Polytechnic Institute, Yangzhou 225000, China

†Department of Transportation Engineering, Yangzhou Polytechnic Institute, Yangzhou 225000, China

‡Suzhou Sip Ourlead Industrial Equipment Co., Ltd, Suzhou 215000, China

§School of economics and management, Sichuan Tourism University, Chengdu 610100, China (Corresponding author, Yaotong2011tony@163.com)

The research aims to reduce the power consumption during URT operation and alleviate urban congestion. In addition, the new era requires high-quality and sustainable development, and URT must focus on energy conservation and emission reduction. The innovation points of the research mainly include the following two points: firstly, efficient utilization of regenerative braking energy to optimize the scheduling and energy-saving of multiple trains; secondly, the use of FA algorithm, an emerging swarm intelligence algorithm, to design efficient attraction models. The research structure is mainly divided into four parts. The first part is a review of relevant research results; The second part is to introduce regenerative braking energy into the optimization of multi train operation scheduling and establish the FURTOSO algorithm; The third part is the validation of the effectiveness and practicality of the proposed research methods; The final part is a summary of the research. This study aims to reduce the energy consumption required for URT traction power supply and achieve the goal of energy-saving optimization for the entire URT line, which is of great significance for the long-term development and environmental protection of URT in the future.

**2. Related Work.** Most of the energy consumed by URT operation is used for traction conditions. Therefore, the research on scheduling optimization of traction energy consumption has extremely important practical significance. The factors affecting traction energy consumption mainly include infrastructure and TO. Currently, most cities have completed the basic construction of URT. However, it is very difficult and costly to reconstruct the infrastructure of the completed lines. Therefore, many related researchers have conducted discussions about the scheduling URT TO. Zhao S et al., in order to adapt to the asymmetric tidal time-varying characteristics of URT line passenger demand and reduce the cost of train operation, studied the introduction of cycle balance into the top-level optimization, and built a low cost oriented integer linear programming model. The results of a case study based on even a certain URT route show that this method not only meets the service level of travel needs, but also improves the efficiency of circulation and utilization [5]. Li X et al. studied and analyzed the overall travel characteristics of passengers after the URT operation in order to study the passengers' choice of URT and public transportation after the URT operation, and then used the random forest algorithm to establish the passenger travel mode selection model after the URT is put into use. The experimental results indicate that travel cost is the most critical factor affecting passengers' travel decisions, and whether to travel during peak hours has a relatively small impact on passengers' choices, which provides support for transportation decisions [6]. Li W et al. proposed a multi-objective optimization model for urban railways to optimize and improve the utilization rate of regenerative braking energy and reduce energy consumption, while ensuring the service quality of the URT system while meeting passenger service needs without increasing the deviation of train running time for one lap. The model was solved using a non dominated sorting genetic algorithm - II. The research results show that the optimal energy-saving schedule reduces total energy consumption by 8.72%, but the deviation of one week train operation time increases by 728 seconds; The total energy consumption decreased by 6.09%, but the deviation of train operation time for one week did not increase [7]. Tang J et al. proposed a data-driven URT schedule optimization method to reduce the operating costs and improve service quality of URT company. They constructed a dual objective optimization model with the goal of minimizing the total waiting time of passengers and the company's departure time. The study conducted an experiment on a certain line in Beijing, and the results verified the effectiveness of this method. This method can provide high-quality and reasonable timetable solutions for URT system managers [8]. URT trains will continuously generate energy conversion during operation. Currently, the energy processing method generated under braking conditions is energy feedback. Then, the RBE is fed back to the AC power grid through a grid connected inverter. If other vehicles are in the traction phase, RBE can be preferentially used. Besides, the attraction model of FA algorithm determines the running time of the algorithm, which can be well used for complex problems such as high-dimensional multi-objective user dynamic optimization. Shen X et al. proposed a new method about URT RBE to enhance the utilization rate of train RBE. The simulation illustrates that the method could enhance the utilization of RBE [9]. Yang Z et al. designed an energy management strategy (EMS) according to deep reinforcement learning to optimize the EMS of URT supercapacitor energy storage systems. This research demonstrates that this method can dynamically adjust the voltage threshold, and significantly improve the energy saving and voltage stabilizing effects [10]. Zhang C et al. proposed a joint optimization model for dual track railway networks to optimize train scheduling in the railway network. Then, this study uses the heuristic algorithm of Lagrange relaxation method to solve the above model. The calculation verifies

the effectiveness and efficiency of the model budget method [11]. Christy J J et al. studied the use of adaptive discrete firefly algorithms to solve broadcast scheduling problems in order to achieve optimal broadcast scheduling algorithms for intelligent transportation systems. This experiment verifies the effectiveness of this method [12]. Based on the above research, it can be concluded that the research on scheduling optimization of train passengers or time is relatively mature. However, there is very little research on multi train coordination and how to coordinate the utilization of RBE and passenger comfort. For enhancing the operational efficiency of URT trains while ensuring passenger satisfaction, the study starts from the perspective of driving strategies and timetables. Then, this study establishes the FURTOSO algorithm based on the FA algorithm to optimize the coordinated operation scheduling of multiple trains.

### 3. Optimization of UTR Train Operation Scheduling Based on FA Algorithm.

**3.1. Construction of Train Dynamics Model and Energy Consumption Model.** Currently, most cities have completed the construction of URTs, but there are problems with the difficulty and high cost of basic design transformation for the completed station routes. Therefore, this study optimizes scheduling and energy conservation from the TO. The operation of URT trains is divided into four working conditions according to the force conditions: traction, cruise, coasting, and braking. When a train is running, it will switch under four states according to line conditions such as speed limits [13, 14]. The initial state of TO must be traction condition. When the train operates under this working condition, it generally adopts the maximum traction force to accelerate, and the braking force is 0. At this time, the train accelerates at the maximum acceleration. The cruise condition is that the time from the start of the train to the braking stop is always affected by the basic resistance; And the train maintains a state of uniform motion. The distance between stations in urban rail transit is relatively short, so cruise mode is usually not used during TO. During TO, the four-stage operation mode is only used in sections with long distance between stations [15, 16]. To reduce the energy consumption required for TO, the coasting condition will be used for transition before entering braking, at which time the train is only influenced by resistance. When a train arrives at a station and needs to stop, braking is generally used, and the train speed continues to decrease. For tracking TO, it is necessary to consider the safety of the front and rear trains during operation. The study only considers increasing the speed limit for tracking trains based on the line speed limit. The calculation is shown in Equation 3.1.

$$V_L = \min \left\{ V_{TSL}, \sqrt{2LA_E} \right\} \quad (3.1)$$

$V_{TSL}$  in Equation 3.1, denotes the speed limit of the tracking train line,  $km/h$ ; is the distance between the tracking train and the vehicle in front,  $m$ ;  $A_E$  denotes the maximum braking speed of the tracking train,  $m/s^2$ . The train dynamics model during URT TO mainly focuses on longitudinal dynamic effects. This study establishes a train dynamics model in view of the above basic principles. The train dynamics model established in the study is considered as a single particle model, and only the train length is considered when calculating the slope. Urban rail trains are usually divided into motor cars and trailers. The motor train is responsible for converting electrical energy into mechanical energy to provide running power for the train; The trailer itself does not have a power unit, so the traction body required for the train is provided by the traction motor of the motor car. The formation mechanism of locomotive traction force and the characteristic curve of train traction are shown in Figure 3.1. In Figure 3.1(a), the vehicle exerts gravity on the rail, and the wheel receives a reaction force from the rail through the contact point. When the traction motor operates, it outputs a torque to the wheels. Which causes the wheel to rotate at the center of the circle, generating a force and a reaction force ; Where is the radius of the circle. At this point, will prevent the sliding action between the wheel and the rail, and will be converted into a force to push the wheel to roll. Figure 3.1 (b) showcases the actual output traction force of a train with different maximum traction forces at different speeds. It is shown in Equation 3.2.

$$F = \zeta F_{\max} \quad (3.2)$$

In Equation 3.2,  $\zeta$  denotes the percentage of the actual output traction acceleration and maximum acceleration; denotes the maximum traction force. The maximum braking force of a train also changes with the speed, and its actual output braking force is calculated as shown in Equation 3.3.

$$A = \Psi A_{\max} \quad (3.3)$$

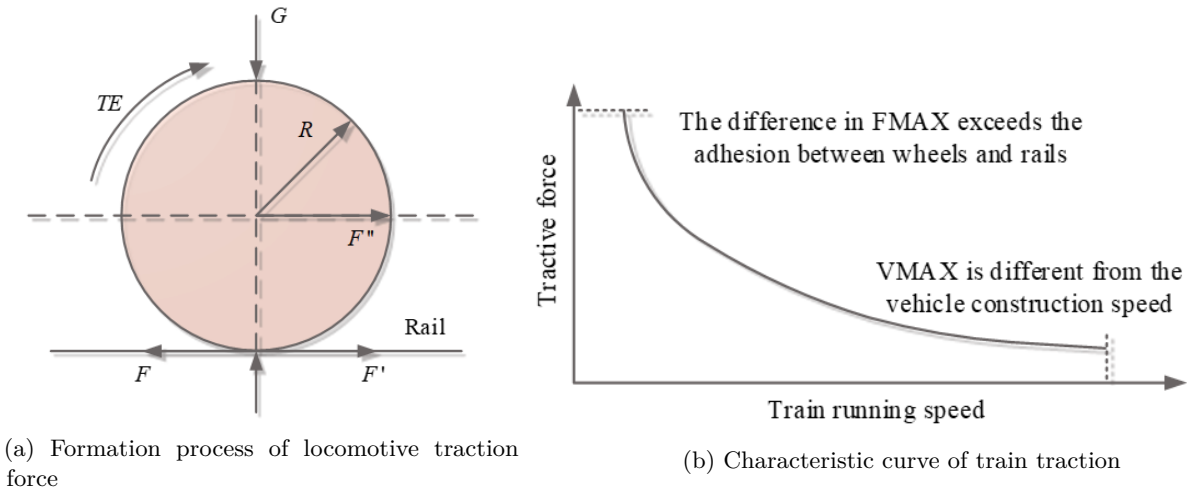


Fig. 3.1: Formation Mechanism of Locomotive Traction Force and Train Traction Characteristic Curve

$\Psi$  in Equation (3) denotes the percentage of the actual output braking acceleration to the maximum acceleration; denotes the maximum braking force. The resistance during TO is not controlled artificially and includes basic resistance and additional resistance based on their causes. The basic resistance is the resistance that a train will encounter during any operation process. Affected by many factors, it is difficult to analyze each factor in a specific way. Therefore, the empirical formula obtained from multiple experiments is utilized for calculating the basic resistance  $B_0$ , as shown in Equation 3.4.

$$B_0 = a + bv + dv^2 \tag{3.4}$$

In Equation 3.4,  $v$  is the running speed;  $a, b, d$  are empirical coefficients. Additional resistance is the resistance generated when passing a special route, which is only related to the environment. The additional resistance is divided into slope additional resistance and curve additional resistance. The calculation of unit slope additional resistance is shown in Equation 3.5.

$$w_i = \frac{W_i}{Mg} \cdot 1000 = 1000 \tan \theta \tag{3.5}$$

In Equation (5),  $W_i$  is the additional resistance of the ramp;  $M$  is the weight;  $g$  is the gravity's acceleration;  $\theta$  is the included angle of the ramp;  $\tan \theta$  will be given during line design. The calculation of the additional resistance  $w_c$  per unit curve is shown in Equation 3.6.

$$w_c = \frac{600}{R} \tag{3.6}$$

After completing the above train dynamics model, it is also necessary to model the train energy consumption. This study aims to achieve scheduling optimization and energy conservation by planning the driving strategy and timing of trains. Therefore, the main consideration is the traction energy consumption during TO, ignoring the equipment energy consumption. The relevant schematic diagram is showcased in Figure 3.2. Figure 2 indicates that the energy required for TO is generated by the substation, and the generated DC power is input into the traction substation. Then, the DC power is transmitted to the conversion device through the train pantograph, converting the DC power into AC power for the generator to operate. Currently, there are two methods for solving the traction energy consumption of a single train: traction work and active current. Calculating the energy consumed by the train through traction work is shown in Equation 3.7.

$$E_1 = \frac{\sum_{i=1}^n Fv \cdot \Delta l_i}{3600} \tag{3.7}$$

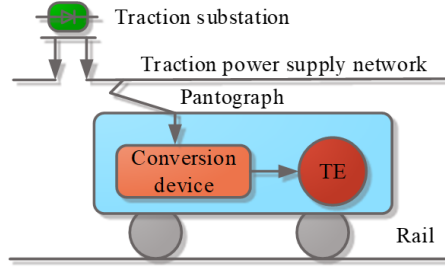


Fig. 3.2: Schematic Diagram of Energy Flow of Traction System

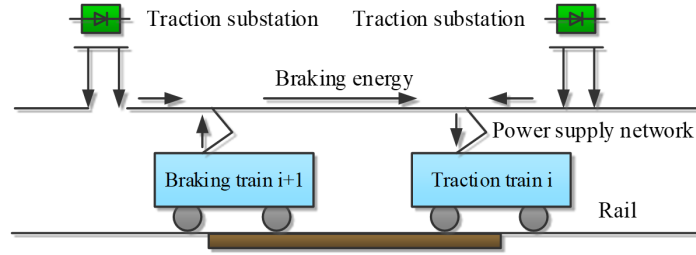


Fig. 3.3: Schematic Diagram of Renewable Energy Utilization

In Equation 3.7,  $F_v$  is the corresponding traction force when the train speed is  $v, kN$ ;  $s$  is the distance the train runs in a section,  $m$ . The energy consumption calculated by active current is shown in Equation 3.8.

$$E_2 = \frac{U_p \{ \sum_{i=1}^n \} [(\bar{I})_t] \Delta t_i}{60 \cdot 100} + \frac{U_p I_{SUE} \sum_{j=1}^m \Delta T_j}{60 \cdot 100} \quad (3.8)$$

$U_p$  in Equation 3.8 is the network voltage of the pantograph,  $V$ ;  $\bar{I}$  serves as the average active current,  $A$ ;  $\Delta t_i$  and  $\Delta t_j$  are respectively the running time under traction and other working conditions,  $\text{min}$ ;  $I_{SUE}$  is the active power of self-use electricity,  $A$ . The active current method is suitable for situations where the energy consumption curve for TO is known. Suitable calculation methods can be used according to different environments in practical applications.

### 3.2. Design of Optimal Algorithm for Urban Rail Transit Operation Scheduling Based on FA.

Compared with railways, URT trains have the characteristics of strong departure periodicity, more stops, and frequent traction and braking switching. Therefore, the energy generated during train braking is extremely high [17]. Traditional air braking methods can cause huge energy losses. Therefore, with the continuous updating of braking technology, URT currently mainly uses electric braking. Electric braking includes regenerative braking (RB) and resistance braking. The RB method can reuse the energy lost by traditional methods and provide it to other trains in the same power supply section for use; If the braking force is insufficient, resistance braking can be used. Due to the high cost of energy storage equipment and its difficulty in promotion, research is mainly aimed at optimizing TO schedules to increase the overlap time between braking and traction of front and rear vehicles. The study then maximizes the RBE from trains and feeds it back to the power supply network for use by other trains, thereby reducing energy losses. Figure 3.3 demonstrates the renewable energy utilization specifically.

In Figure 3.3, this study assumes that train  $i + 1$  is in the inbound braking condition, and the generated RBE is transmitted to the power supply network in the form of electrical energy. Meanwhile, train leaves the station under traction condition. This can use the RBE  $E_r$  generated during the overlapping time of two train working conditions. In practical situations, it is not possible to ensure the complete overlap of the working

conditions of two trains. Therefore, the utilization rate of RBE generated by train  $i + 1$  can be calculated based on the train  $i$ 's overlapping time. The utilized RBE  $E_{bu}$  is Equation 3.9.

$$E_{bu} = \frac{Rr \cdot t_0}{t_a} \quad (3.9)$$

$t_0$  in Equation 3.9 is the overlapping time of two train working conditions;  $t_a$  serves as the braking time of train  $i + 1$ . After the above preparation, it can be concluded that the URT TO scheduling optimization problem is a strategic problem that extends from two stations to multiple stations. The operation process between a single train and two stations is a known distance between two stations. There are different line conditions such as different ramps and curves on the section line, and the process of trains stopping from the starting station to the terminal station. From this, it can be concluded that the URT train scheduling optimization problem is essentially a linear problem. Its description is based on line speed limits, operation time, and station spacing as constraints, with the goal of minimizing TO energy consumption. Then it optimizes the running schedule based on the FA algorithm and designs the FURTOSO algorithm. There are four working conditions during TO, and understanding the dynamic relationship of trains in different conditions can be the basis for constructing the FURTOSO algorithm. Therefore, in view of Newton's second law, the dynamic equations under four operating conditions can be obtained. Then, the study uses the FA algorithm to randomly set several different firefly locations within the solution interval. Subsequently, it iteratively updates these solutions and ultimately finds the theoretically optimal solution. Compared with other intelligent optimization algorithms, FA algorithm has the advantages of fewer parameters, good stability, and easy operation. The core idea is that fireflies with lower absolute brightness congregate with fireflies with higher absolute brightness. The objective function is defined by the absolute brightness of the firefly. In this study, the objective function values corresponding to the absolute brightness  $I_p$ 's positions and  $\vec{z}_p$  of the firefly  $p$  at the  $\vec{z}_p = (\vec{z}_{p1}, \vec{z}_{p2}, \dots, \vec{z}_{px})$  position are equal. The relative brightness of firefly  $p$  versus  $q$  firefly is defined as Equation 3.10.

$$I_{pq}(l_{pq}) = I_p e^{-\vartheta l_{pq}^2} \quad (3.10)$$

$I_{pq}$  in Equation 3.10 is the distance between fireflies;  $\vartheta$  is the light absorption coefficient. The greater the relative measure, the greater the attraction. The calculation of attraction is shown in Equation 3.11.

$$AF_{pq}(d_{pq}) = AF_0 e^{-\vartheta d_{pq}^2} \quad (3.11)$$

In Equation 3.11 is the Cartesian distance between fireflies; is the greatest attraction. Then, the position update of the firefly after moving to can be obtained, as shown in Equation 3.12.

$$\vec{z}_q(t+1) = \vec{z}_q(t) + AF_{pq}(l_{pq})(\vec{z}_p(t) - \vec{z}_q(t)) + \tau \vec{\mu}_q \quad (3.12)$$

$\vec{z}_p(t)$  and  $\vec{z}_q(t)$  in Equation 3.12 are the spatial locations of fireflies  $p$  and  $q$ , respectively;  $t$  is the number of iterations;  $\tau$  is the step length factor;  $\vec{\mu}_q$  denotes a random number vector. To sum up, the process of FA algorithm can be obtained. Figure 3.4 demonstrates the details.

In Figure 3.2, the process first initializes the algorithm, and then moves the firefly. It then updates the absolute metric and iterates through the loop. Optimization of single TO scheduling mainly involves allocating the time for each operating condition. Optimization of multi TO scheduling is achieved by adjusting the train schedule. In short, multi station optimization of a single train is an extreme point problem; Multi train multi station optimization is a balancing problem. The FURTOSO algorithm is based on the operation optimization control of a single train to further optimize the train schedule. Constraints for multi TO include vehicle parameters, line conditions, and train schedules. The train timetable ensures that the train can work securely, specifying the parameters required for operation. Therefore, designing a reasonable timetable can improve the overlapping time of braking and traction for different trains on the same line to decrease energy consumption. The details are shown in Figure 3.4.

Figure 3.5 illustrates the v-t curve of trains 1 and 2 running on the same line. The TO's overlapping time can be changed by changing the departure interval, operation time between stations, and dwell time. Therefore,



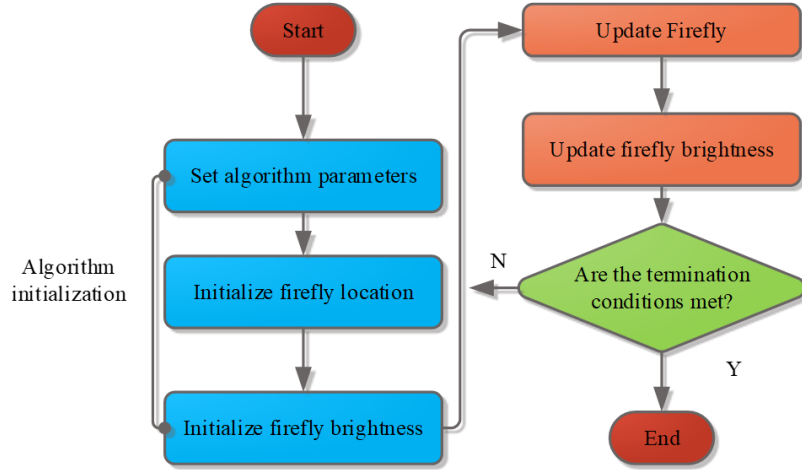


Fig. 3.4: Flow of FA Algorithm

reasonable scheduling has practical significance for URT TO scheduling optimization. For multi TO scheduling optimization problems, it is necessary to find the interval  $H = \{h_1, h_2, \dots, h_{n-1}\}$  with the lowest TEC for all trains with a number of  $n$ . The study assumes that the time when the  $i^{th}$  train departs from the  $j^{th}$  platform is  $Y_i^j$ , the time when it arrives at the  $j+1^{th}$  platform is  $D_i^j$ , and the departure interval between the  $i^{th}$  and  $i+1^{th}$  trains at the starting platform is  $h_i = D_{i+1}^1 - D_i^1$ . The main purpose is to minimize the energy consumption of all platforms through which the train passes and generate the most regenerative energy during operation. Therefore, based on the constraints of acceleration, traction, and braking force, an optimization model for multi train and multi platform operation scheduling is established. Model 1 represented by Equation 3.13.

$$\begin{cases} \min E_1 = \sum_{i=1}^n \sum_{j=1}^{k-1} \int_{D_i^j}^{Y_i^{j+1}} F[v(t)] dt \\ \left\{ \begin{array}{l} K_{min} \leq D_i^j - Y_i^j \leq K_{max}, i = 1, 2, 3, \dots, n; j = 2, 3, \dots, k-1 \\ \sum_{i=1}^{k-1} D_{i+1}^1 = T_0 \\ H_{min} \leq D_{i+1}^1 - D_i^1 \leq H_{max} \\ \sum_{j=1}^{k-1} (Y_i^{j+1} - Y_i^j) \\ v(t) \leq \min(V_{max}, \sqrt{2LA_g}) \\ |Y(t)| \leq Y_{max} \\ F[v(t)] = \zeta F_{max}, \zeta \in [0, 1] \\ A[v(t)] \leq A_{max} \\ v(0) = v(D_i^j) = v(Y_i^j) = 0, i = 1, 2, \dots, n \end{array} \right. \\ \max E_2 = \sum_{i=1}^{n-1} \sum_{j=2}^{k-1} \int_{Y_{j+1}^j}^{Y_{i+1}^j} A[v(t)] v * 95\% * t_{io1}^j / t_{i+1bk}^j dt \end{cases} \quad (3.13)$$

The  $E_1$  and  $E_2$  of Equation 3.13 are the TEC and the total energy utilized for regenerative energy of the train during operation time;  $T_o$  is the departure time interval between the first train and the last train;  $t_{io1}^j$  is the overlapping time period between the braking time of the  $i+1$ -th train and the acceleration time of the  $i$ -th train at the  $j$ -th station;  $t_{i+1bk}^j$  is the braking time period of the  $i$ -th vehicle at the  $J$  station;  $k_{min}$  and  $k_{max}$  are the minimum and maximum dwell times;  $H_{min}$  and  $H_{max}$  are the minimum and maximum intervals, respectively. To find the interval with the lowest TEC during the operation of all trains, a model with  $h_i$  as an indirect variable is studied and equivalently converted to a model with  $h_i$  as a decision variable, as shown in

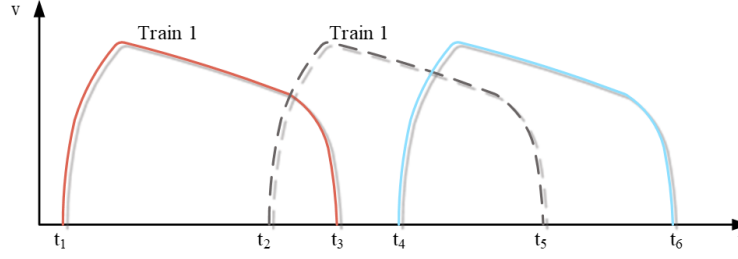


Fig. 3.5: Operation Curve of Multiple Trains at the Same Time

Equation 3.14.

$$\left\{ \begin{array}{l} \min E_1 = \sum_{i=1}^n \sum_{j=1}^{k-1} \int_0^{T_i^j} F[v(t)] v(t) dt \\ \left\{ \begin{array}{l} K_{min} \leq D_i^j \leq K_{max}, i = 1, 2, 3, \dots, n; j = 2, 3, \dots, k-1 \\ \sum_{i=1}^{k-1} h_i = T_0 \\ H_{min} \leq h_i \leq H_{max}, i = 1, 2, \dots, n-1 \\ \sum_{j=1}^{k-1} (D_i^j - T_i^j) = T \\ v(t) \leq \min(V_{max}, \sqrt{2LA_g}) \\ |Y(t)| \leq Y_{max} \\ F[v(t)] = \zeta F_{max}, \zeta \in [0, 1] \\ A[v(t)] \leq A_{max} \\ v(0) = v(T_i^j) = v(T_i^j + v_i^j) = 0, i = 1, 2, \dots, n; j = 1, 2, \dots, k-1 \end{array} \right. \\ \max E_2 = \sum_{i=1}^{n-1} \sum_{j=2}^{k-1} \int_0^{T_{i+1}^j} A[v(t)] v * 95\% * t_{i01}^j / t_{i+1bk}^j dt \end{array} \right. \quad (3.14)$$

$T_i^j$  in Equation 3.14 denotes the total operation time of the  $i$  vehicle from Station  $j$  to Station  $j+1$ ;  $h_i$  is the interval between the  $i^{th}$  vehicle and the  $i+1^{th}$  vehicle. To sum up, a multi TO in view of the FURTOSO algorithm can be obtained, as shown in Figure 3.6.

#### 4. Result Analysis of the FURTOSO Algorithm.

**4.1. Performance Analysis of the FURTOSO Algorithm.** For proving the FURTOSO algorithm's function proposed in the study, Matlab software was used for simulation. To more scientifically test the superiority of the FURTOSO algorithm in handling the URT train optimal scheduling problem, comparative experiments were conducted using currently commonly used bacterial foraging optimization (BFO), particle swarm optimization (PSO), and genetic algorithm (GA) [18, 19, 20]. The study ran the four algorithms independently 50 times. Then it also takes the optimal operation result of the objective function value as the final optimization result.

Figure 4.1 showcases the fitness convergence curve results of different algorithms. Figure 4.1 demonstrates that compared to the other three algorithms, the FURTOSO algorithm possesses higher convergence accuracy and faster convergence speed for solving the objective function. The algorithm only needs 76 iterations to reach a stable state, with a fitness value of 0.6827. The BFO algorithm has the worst convergence effect, requiring 815 iterations to reach the objective function. The PSO algorithm requires 135 iterations, and the GA algorithm requires 203 iterations. The FURTOSO algorithm proposed in the study can quickly converge to the target state due to its strong local search ability, which can quickly and easily find the optimal solution in a region. Due to the lack of dynamic speed adjustment, the PSO algorithm is prone to falling into local optima, resulting in lower convergence accuracy and difficulty in convergence. The BFO algorithm mainly focuses on adjusting parameters during operation, so it is difficult to ensure the progress and convergence speed of the solution when

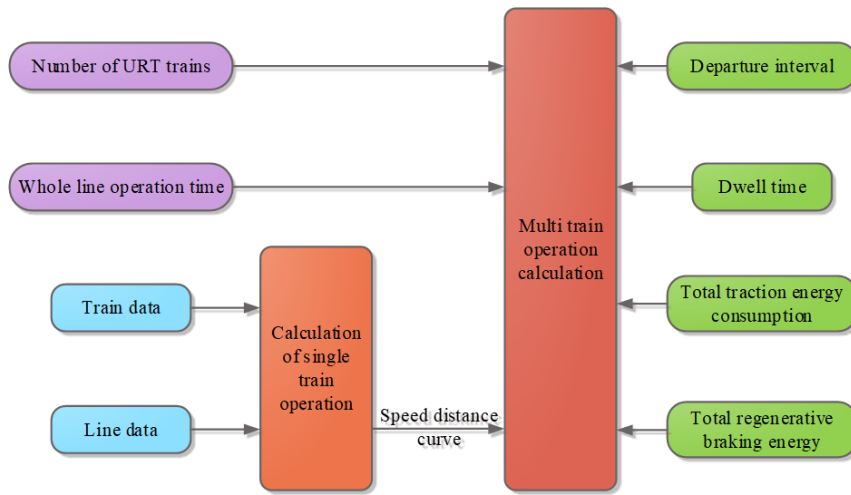


Fig. 3.6: Multi-Train Operation Process Based on FURTOSO Algorithm

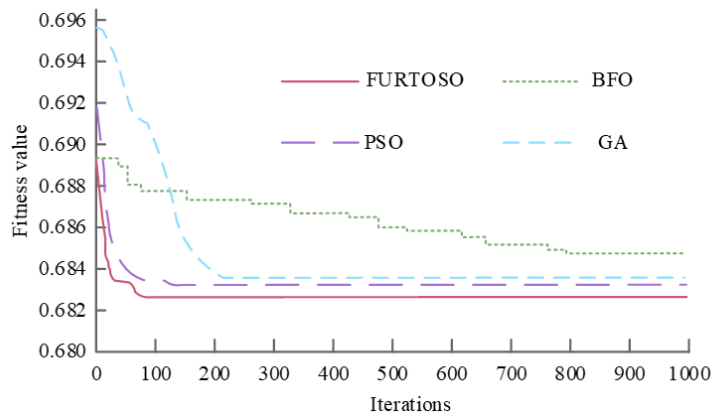


Fig. 4.1: Convergence Change Curve Results of Different Algorithms

optimizing different types of problems. The GA algorithm is prone to issues of non-standard and inaccurate encoding, and its local search ability is poor, resulting in very slow convergence speed.

Table 4.1 indicates the comparison of the four algorithms' optimization. Table 4.1 illustrates that overall, the optimal value (OV) and average OV of the FURTOSO algorithm are the smallest,  $6.8266945e-001$  and  $6.82717688e-001$ , respectively. The optimal objective function value can be obtained by minimizing the number of iterations. The OV and average OV of BFO algorithm are  $6.83921100e-001$  and  $6.84410765e-001$ , respectively. The FURTOSO algorithm proposed in the study has simple mathematical principles, fewer parameters, and minimal impact of parameters on the algorithm. Therefore, it can obtain optimization results in a shorter time, and it can simultaneously achieve high operational efficiency and accurately solve URT scheduling optimization problems. The parameters of the BFO algorithm do not have self-adaptability, so its operational efficiency cannot be guaranteed; The PSO algorithm requires selecting appropriate parameters to achieve optimal results for different problems, which can affect the efficiency and optimization results of the algorithm; The ability of GA algorithm to explore new spaces is limited, and it will consume a lot of time when conducting a large amount of calculations. In summary, the FURTOSO algorithm has better performance. It can efficiently and

Table 4.1: Optimization Results of Different Algorithms

Algorithm	FURTOSO	BFO	PSO	GA
Optimal Value	6.8266e-001	6.8392e-001	6.8335e-001	6.8365e-001
Average OV	6.8271e-001	6.8441e-001	6.8355e-001	6.8374e-001
Tending to OIA	76	815	135	203

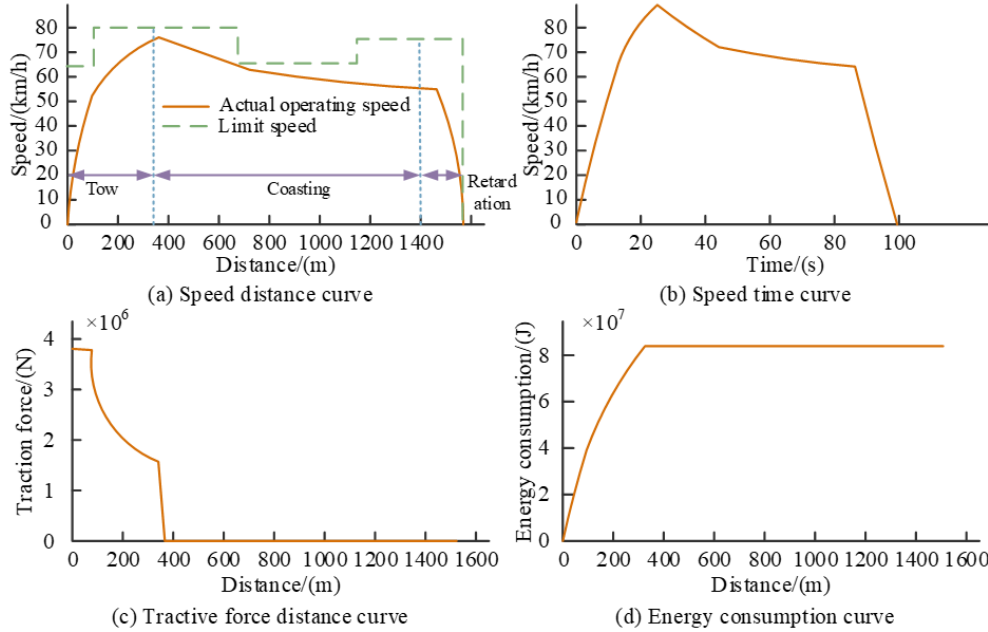


Fig. 4.2: Optimization Results of Single Train and Single Station Based on FURTOSO Algorithm

accurately solve the problem of URT operation scheduling optimization.

**4.2. Application Analysis of FURTOSO Algorithm.** For proving the FURTOSO algorithm's availability in practical applications, the simulation took Chengdu Metro Line 8 as the research object. The weight of a train consists of its own weight and the weight of passengers. It uses six carriages and is organized into four motor cars and two trailers, with 230 people for each trailer and 250 people for each motor car. This study assumes a weight of 60 kg per passenger, resulting in a total train weight of 288.6 t. Other parameters of the train are set as follows, with a length of 120m and a maximum operating speed of 80km/h; The basic resistance parameters  $a$ ,  $c$ , and  $c$  are 2.031, 0.0622, and 0.001807, respectively; The maximum acceleration and deceleration are both 1m/s<sup>2</sup>; The minimum and maximum dwell times are 30s and 45s respectively; The minimum and maximum departure times are 2 min and 11 min respectively; The interval between the first train and the last train is 990 minutes. Chengdu Line 4 has a total length of 28.8km, an operating time of  $52 \pm 0.5$ min, and a maximum operating speed of 80km/h. It starts at Lianhua Station and ends at Shilidian Station, with a total of 24 stations.

Figure 4.2 shows the optimization results of single train and single station operation scheduling based on the FURTOSO algorithm. From Figure 4.2 (a) to (d), the optimization results of distance speed, time speed, distance traction, and distance energy consumption for micro TO are summarized. From the overall analysis of Figure 8, with a fixed total operating time, the train traction time and coasting time are longer, and the cruise time and braking time are shorter. The energy consumption is  $8.42 \times 10^7 J$ , which is 20.04% lower than

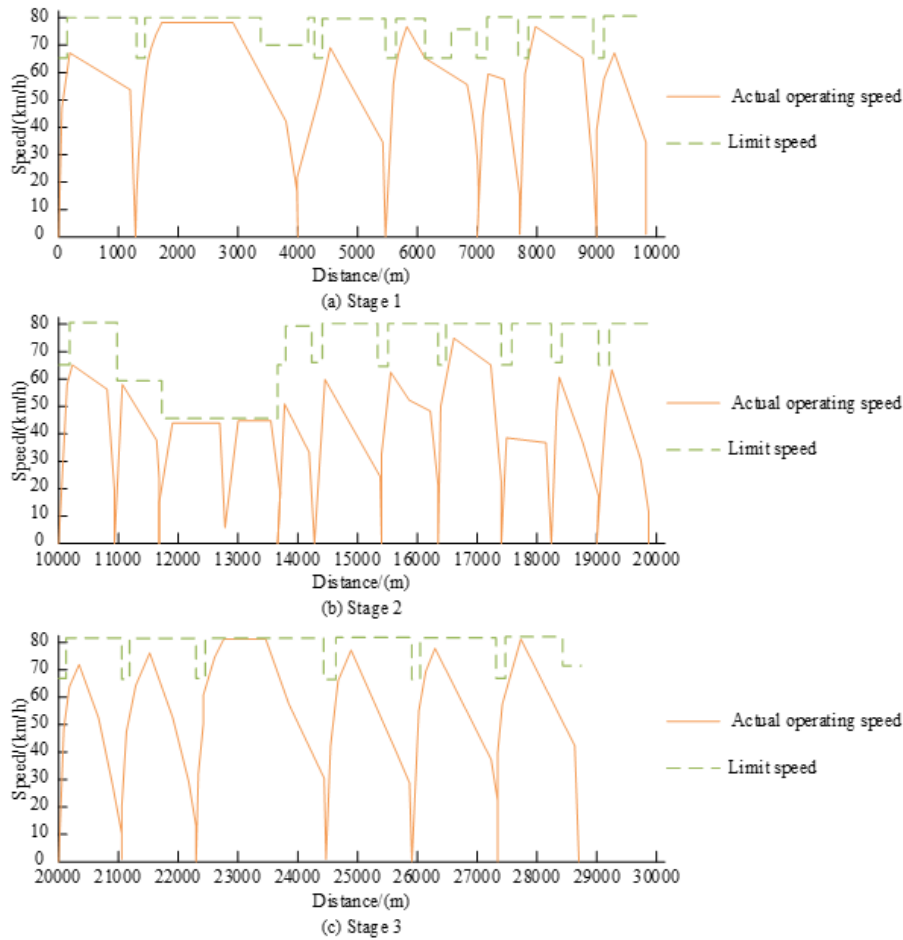


Fig. 4.3: Speed-Distance Curve of a Single Train Passing 24 Stations Based on the FURTOSO Algorithm

the energy consumption of  $10.53 * 107J$  before optimization.

Figure 4.3 (a) to (c) show the speed distance variation curve of a single train passing through 24 stations from Lianhua to Shilidian based on the FURTOSO algorithm, from stage 1 to stage 3. Most trains operate in a three-stage strategy mode of traction, coasting, and braking between stations. In a few stations, it uses a four-stage mode of operation, namely, traction - coasting - cruise - braking. The slope from Jiuxing Avenue to Yongfeng Station is relatively large, reaching 35 %. Therefore, a four-stage model is used for optimization.

The research compares the consumption results of this part before and after optimization, and Table 4.2 demonstrates the results. Table 4.2 showcases that after optimizing Chengdu Metro Line 8 using the FURTOSO algorithm, the energy consumed by trains operating in most station sections has decreased. Due to the time evolution of the FURTOSO algorithm for TO between stations, the overall energy consumption of TO has been reduced by 12.78% from the perspective of the overall line effect. This confirms the feasibility of using the FURTOSO algorithm for URT scheduling optimization. The full day operation time of Chengdu Metro Line 8 is 990 minutes, with a total of 157 trains departing; From this, it can be concluded that the average departure interval is 381 seconds, and the overall operation time of a single train is 3120 seconds. If the station conducts uniform departure, a maximum of 9 vehicles can operate simultaneously on the line. Therefore, the study takes 10 trains as a group to optimize the solution. The optimized set of train departure intervals is (288, 288, 288, 288, 289, 497, 497, 497, 497), with an overlap time of 895s. Based on this, there are about 17 sets of trains in total throughout the day, and the optimization results for multiple trains are obtained, as shown

Table 4.2: Comparison of Results Before and After Train Optimization at Large Gradient Stage

Sec.	Start/End Station	Oper. Time/s	Actual Energy	Opt. Energy (J)	Energy Saving Rate
1	Lianhua-Wenxing	93	7.42	5.21	29.78
2	Wenxing-Jiang'an Campus	207	20.63	12.89	37.52
3	Jiang'an-Pearl River	97	8.31	6.70	19.37
4	Pearl River-Shunfeng	101	10.54	8.45	19.83
5	Shunfeng-Sanyuan	71	4.95	3.87	21.82
6	Sanyuan-Shiyang	85	8.76	6.69	23.69
7	Shiyang-Qing'an	71	5.83	7.32	-25.56
8	Qing'an-Banjialin	75	6.59	5.01	23.98
9	Banjialin-Gaopeng Ave.	197	13.86	14.28	3.03
10	Gaopeng Ave.-Jiuxing Ave.	69	5.23	4.07	21.61
11	Jiuxing Ave.-Yongfeng	62	4.18	3.12	25.36
12	Yongfeng-Fangcaojie	107	7.72	4.49	41.84
13	Fangcaojie-Nijiaqiao	81	6.68	5.16	23.65
14	Nijiaqiao-Wangjiang Campus	76	7.52	6.68	11.17
15	Wangjiang Campus-East Lake	89	2.93	1.62	45.27
16	Donghu Park-Dongguang	87	5.66	4.28	24.38
17	Dongguang-Jingyuansi	86	5.89	6.01	-2.04
18	Jingyuansi-Dongda Road	106	7.57	8.57	-13.21
19	Dongda Road-Shuangqiao Rd.	118	9.12	10.53	-15.46
20	Shuangqiao Rd.-Wannian Rd.	151	15.62	19.12	-22.86
21	Wannian Rd.-Shanbanqiao Rd.	117	10.36	8.71	15.93
22	Shanbanqiao-Dongjiao Memory	108	10.15	8.63	18.42
23	Dongjiao Memory-Shilidian	100	9.93	9.06	8.76
Total	Lianhua-Shilidian	2354	195.45	170.47	12.78

in Table 3. Table 3 shows that train departure intervals are at the peak end of the overlapping time. The total overlap time of the entire line is 15475s, the utilization rate of RBE is 30.1%, and the TEC is  $2.661 * 1011J$ . Compared to the actual energy consumption of  $3.059 * 1011J$ , the optimized energy saving rate of the FURTOSO algorithm is 13.03%. In summary, the FURTOSO algorithm has excellent applicability in the operation and scheduling optimization of URT trains.

**5. Conclusion.** URT has the advantages of green safety, speed and punctuality. It plays a significant role in promoting the modernization process, improving the transportation environment, guiding and optimizing the urban spatial layout, and driving the innovative development of the urban economy. However, the large-scale and high-speed development of URT system in China. Therefore, how to promote the application of advanced technology in the URT industry, optimize the operation scheduling of URT, and reduce more energy consumption is crucial for smart cities. In response to the above problems, a FURTOSO algorithm was established to optimize TO scheduling with the minimum energy consumption of TO as the goal. The experiment showcases that the FURTOSO algorithm has higher convergence accuracy and faster convergence speed for solving the objective function. It only requires 76 iterations to reach a stable state, with a fitness value of 0.6827. The BFO algorithm requires 815 iterations, the PSO algorithm 135 iterations, and the GA algorithm 203 iterations to stabilize in the target state. In practical applications, train departure intervals are all at the peak end of the overlapping time, with a utilization rate of 30.1% of RBE and a TEC of  $2.661 * 1011J$ . Compared to the actual energy consumption of  $3.059 * 1011J$ , the optimized energy saving rate of the FURTOSO algorithm is 13.03%. In summary, the FURTOSO algorithm has good performance and optimization effects. However, there are still shortcomings in the research. When modeling the operation process of multiple trains, only the overlapping time of traction and braking before and after the maximum is considered, without considering the passenger's



- [11] Zhang, C., Gao, Y., Yang, L., Gao, Z. & Qi, J. Joint optimization of train scheduling and maintenance planning in a railway network: A heuristic algorithm using Lagrangian relaxation[J]. *Transportation Research Part B: Methodological*. **134** (2020)
- [12] Christy, J., Rekha, D. & Vijayakuma, V. Carvalho G H S. *Optimal Broadcast Scheduling Method For VANETs: An Adaptive Discrete Firefly Approach*. **39**, 8125-8137 (2020)
- [13] Yldrm, D., Akit, M., Yolaan, C., Pul, T., Ermi, C., Aghdam, B., Adrc, I. & Ermi, M. Full-Scale Physical Simulator of All SiC Traction Motor Drive with Onboard Supercapacitor ESS for Light-Rail Public Transportation. *IEEE Transactions On Industrial Electronics*. **67**, 6290-6301 (2020)
- [14] Mallia, D., Mitchell, L., Kunik, L., Fasoli, B. & Lin, J. Constraining Urban CO<sub>2</sub> Emissions Using Mobile Observations from a Light Rail Public Transit Platform. *Environmental Science And Technology*. **54**, 15613-15621 (2020)
- [15] Teng, H., Toughrai, T., Yu, T. & Ozawa, R. High-Speed Rail Station Interconnectivity and Ridership. *Journal Of Transportation Technologies*. **12**, 193-208 (2022)
- [16] Zdrail, L., Kalytchuk, S., Hola, K., Petr, M. & Zboril, R. carbon dot-based tandem luminescent solar concentrator. *Nanoscale*. **12**, 6664-6672 (2020)
- [17] Kim, K., Park, K. & Nelson, A. Impacts of Light Rail Transit on Labor Participation and Housing Affordability in the U. S.: *Longitudinal Analysis Using Propensity Score Matching*. **2675**, 419-431 (2021)
- [18] Sarkar, D. & Banerjee, S. BFO-based firefly algorithm for multi-objective optimal allocation of generation by integrating renewable energy sources. *International Journal Of Grid And Utility Computing*. **12**, 67-80 (2021)
- [19] Li, T., Kabanikhin, S., Nakamura, G., Wang, F. & Xu, D. An inverse problem of triple-thickness parameters determination for thermal protective clothing with Stephan-Boltzmann interface conditions. *Journal Of Inverse And Ill-posed Problems*. **28**, 411-424 (2020)
- [20] Gao, F., Zhang, Q., Han, Z. & Yang, Y. Evolution test by improved genetic algorithm with application to performance limit evaluation of automatic parallel parking system. *IET Intelligent Transport Systems*. **15**, 754-764 (2021)

*Edited by:* Zhenling Liu

*Special Issue on:* Cloud Computing for Intelligent Traffic Management and Control

*Received:* Apr 21, 2023

*Accepted:* Aug 18, 2023





## METHOD FOR IDENTIFYING MOTOR VEHICLE TRAFFIC VIOLATIONS BASED ON IMPROVED YOLOV NETWORK

**Abstract.** The use of traditional manual supervision means to deal with motor vehicle traffic safety violations can result in a large amount of wasted manpower and oversight problems. To assist road managers in better directing traffic order and managing traffic situations, the study proposes an improved target tracking network model. Simple online real-time tracking, deep correlation metrics, and cascading open-source computer vision libraries are combined to create a tracking model for motor vehicle traffic infraction recognition. Pursuant to the experimental findings, the Institute's upgraded target recognition network model had accuracy and recall rates of 95.7% and 99.7%, respectively, with an accuracy rate of 16.6% higher than the model's historical counterpart. The recognition accuracy of the constructed motor vehicle traffic violation recognition and tracking model regarding the three basic traffic violations was 98.2%, 98.7%, and 97.9%, respectively; the missed detection rate was 2.0%, 0.31%, and 2.1%, respectively; and the false detection rate was 0.17%, 0.31%, and 0%, respectively. It shows that the improved network model of the study is advanced and the motor vehicle traffic offence model has a good recognition rate and stable performance, which can assist traffic managers in their operations to a certain extent.

**Key words:** YOLOV; motor vehicles; SE; CBAM; traffic offences; OpenCV

**1. Introduction.** Motor vehicles have brought great convenience to the travel of human life, but the corresponding number of motor vehicle traffic accidents is also rising year by year. The use of artificial intelligence technology to build an intelligent traffic management system nowadays has become a hot spot for research [1, 2]. Intelligent traffic systems mainly cover traffic monitoring systems, vehicle management systems and advanced traffic management systems, etc. Among them, traffic monitoring systems can assist traffic managers to deal with various traffic accidents effectively and quickly. Traditional traffic video surveillance systems generally use a large number of human assistances, the workload involved is large, but also more prone to errors [3]. First, motor vehicles have various behaviors on the road, and the complexity of the traffic scene makes the task of traffic violation identification extremely challenging. Secondly, the traditional traffic violation identification methods will encounter some limitations when dealing with large-scale data. For large-scale traffic video data, traditional methods based on manually design features face the difficulty of feature extraction and high computational complexity. Therefore, a more optimized and efficient approach to address these challenges needs to be sought. Finally, the existing traffic violation identification methods still have some limitations in terms of accuracy and robustness. Due to the complexity of traffic scenarios and changes in video data, existing methods may in some cases misjudgment or fail to accurately distinguish between illegal and non-illegal behaviors. Therefore, the improved methods need to better solve these problems and improve the accuracy and stability of the identification methods. By optimizing the network structure and feature extraction methods, we hope to address the challenges faced by existing methods in terms of traffic scene complexity, large-scale data processing, and accuracy. The main objective of this study is to provide a more efficient, accurate, and robust method for identifying traffic violations, thereby helping to improve traffic supervision and road safety levels, improve the smoothness of road traffic, and thus save people's time and resources. And then formulate more reasonable Transportation planning and route optimization strategies to improve the efficiency of the overall transportation system. Reduce accident risk and provide important support for decision-making. The study is divided into four parts. The first part discusses the results of relevant research in recent years, the second part focuses on the construction of an improved motor vehicle model, the third part validates the performance of the constructed model and the fourth part draws conclusions from the study.

**2. Related Works.** Many scholars have conducted in-depth research in intelligent systems put into intelligent management of motor vehicles. Motor vehicle traffic violations may lead to traffic accidents, road congestion and unequal traffic distribution. Therefore, accurate identifying and monitoring of motor vehicle

traffic violations is critical to improving road safety and traffic efficiency. This paper improves the YOLOV network to perform better in identifying motor vehicle traffic violations. Through this study, computer vision and deep learning techniques can be better understood and applied to solve traffic safety problems. This study is of great significance for traffic management departments, traffic safety researchers, and technical personnel in related fields. By developing accurate and efficient methods for identifying motor vehicle traffic violations, the efficiency of traffic enforcement can be improved and provide more data support for traffic management, so as to help decision makers to develop more reasonable traffic policies and planning. Sungan et al. in order to propose a CNN-based object detection model that incorporates a TDM-trained multimodal YOLO framework of models to manage the road pothole problem [4]. This method can be used for the analysis of motor vehicle traffic violation recognition, such as the recall and accuracy of the recognition model. Using wireless sensor networks and RFID technologies, Changhong et al. created a framework for an intelligent logistics supply chain management system to increase location accuracy and coverage [5], This framework can be used to establish an identification framework for this study. Fraser et al. artificially analyzed the factors that trigger motor vehicle unsafe events, used cross-case studies to compare potential hazard triggers, and experimentally validated the effectiveness of the proposed study [7]. Zhang et al. created a four-scale detection structure for the purpose to increase the recognition accuracy of long-range tiny target detection [8]. To further enhance the accuracy of YOLO vehicle categorization, Azimjonov et al. suggested a bounding box-based vehicle tracking method and validated the effectiveness of its experimental algorithm [9]. Ahmed et al. presented a fusion transfer learning technique in a deep learning tracking framework to boost the detection model's precision, and they empirically demonstrated its efficacy [10]. In an effort to reduce tracking losses caused by processing occlusion and to achieve more stable tracking, Duan C et al. used a depth-ordered multi-target tracking algorithm to filter multi-target tracking of pedestrians and vehicles in traffic scenes [11]. An experimental study showed that the proposed algorithm can reduce ID switching in real-time traffic. Huiyuan et al. used asymmetric convolution with depth-separable convolution to reduce model parameters to speed up model recognition tracking by gathering vehicle position, trajectory, and other important driving parameters to meet the demands of autonomous driving in a timely manner, and they validated the efficacy of the proposed multi-vehicle tracking framework through in-depth experiments [12]. To address the issues of unsatisfactory detection rate in the YOLOv5 structure, Dong et al. suggested a kind of algorithm that introduces C3Ghost in the YOLOv5 neck structure [13]. Comparison experiments with the YOLOv5 structure were used to demonstrate the effectiveness of their improved algorithm. Wei et al. used the improved YOLOv5 detection to identify riders and helmets and validate their algorithm performance in subsequent experiments [14]. An enhanced YOLO fusion joint-non-maximum suppression model was put forth by Tan et al. to increase the precision of small target identification. The performance of their model was verified through comparison experiments with four existing detection methods [15].

In summary, when conducting further comparative analysis on the use of multi-objective recognition network models in the transportation field, the first thing to consider is the achievements and potential they have achieved. This network model has shown its ability to efficiently handle complex traffic scenarios, such as identifying oncoming vehicles, pedestrians, and even determining the color of traffic lights. By adding appropriate algorithms, such as deep learning or neural network algorithms, the recognition and prediction efficiency can be further improved. However, there are few relevant research materials in the field of motor vehicle Moving violation identification. On this basis, a method for developing a basic model for identifying and tracking automobile traffic safety violations is proposed, which combines deep correlation measurement, open-source computer vision library, and simple online real-time tracking algorithms.

### 3. Model Construction for Identification and Detection of Motor Vehicle Traffic Violations.

The multi-objective tracking model is enhanced in this study. The improved model is further fused within open-source computer vision library to construct a motor vehicle traffic violation recognition pursuit model.

#### 3.1. Construction of a Motor Vehicle Tracking Recognition Model based on YOLOv5s.

YOLOv5 It has a faster reasoning speed while maintaining its accuracy. This is achieved by employing some optimization techniques, such as feature pyramid networks, adaptive receptive fields, and cross-stage feature fusion. YOLOv5 Some improvements in the network architecture, using CSPDarknet53 as the backbone network to improve the accuracy of detection, through techniques such as feature aggregation and attention mechanism.

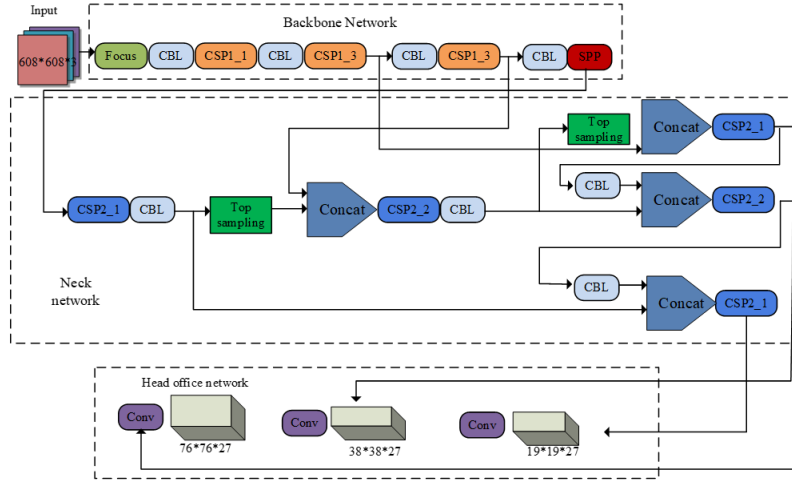


Fig. 3.1: YOLOv5s network structure schematic

Compared to other target detection models, YOLOv5 has smaller models and occupies less hard disk memory and video memory resources, which allows it to operate in more resource-constrained environments. YOLOv5 provides a complete set of code base and pre-trained models that can be easily used and deployed. In addition, it supports multiple platforms and frameworks, including PyTorch, TensorFlow, etc., and therefore is more flexible and scalable. You Only Look Once (YOLO) is primarily used for the detection of multiple target categories and locations. YOLO does not only manage to identify the type of these objects, but also to mark their location. Because it only needs to look once, YOLO is known as a Region Free (RF) method, and in contrast to region-based (RB) methods, YOLO does not need to find the area where the target is likely to be in advance. As the network model of the YOLO family of algorithms has fewer parameters and runs faster, making it more convenient for experiments on embedded platforms, the study builds a corresponding target detection model. YOLOv5 is the OneStage family of algorithms, and its confidence level is given in equation 3.1.

$$Confidence = Pr(Object) \times IOU_{pred}^{truth} \quad (3.1)$$

In equation 3.1,  $P$  represents the likelihood of containing objects within the bounding box and indicates the degree of accuracy of the prediction using the loss function IOU. YOLOv5 will provide different detection models for different object size detection, such as YOLOv5s, YOLOv5m, YOLOv5l, YOLOv5x. The YOLOv5s network structure is demonstrated in Figure 3.1 [15].

In Figure 3.1, the model is divided into four structures. The input rated data is sliced and processed, and the image pixel values are extracted and then cut to enhance the perceptual field and reduce image information loss. The above data undergoes a convolution operation and then enters the CSP layer, The research introduces the Attention Mechanism, which weights and aggregates all historical information to reduce important information loss. The additive model is used to calculate attention scores, as shown in equation 3.2.

$$s(v'_i, v'_t) = \tilde{v}_T \tanh(W_1 v'_i + W_2 v'_t) \quad (3.2)$$

In equation 3.2,  $v'_i$  represents the encoding of historical sequence information after dimensionality reduction,  $v'_t$  show  $t$  encoding of input information after time dimensionality reduction.  $W, v$ , these are the network parameters of the Department of Science. Focus state  $s_t$  expressed as  $H$ , the weighted sum of is expressed in equation 3.3.

$$s_t = \sum_{j=1}^t \alpha_j h_j \quad (3.3)$$

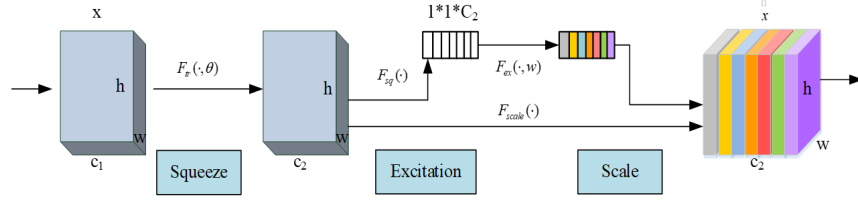


Fig. 3.2: SE schematic diagram

In equation 3.3, combine  $s_t$  and  $h_t$ , the probability calculation expression for motor vehicle identification and monitoring is shown in equation 3.4.

$$y_t = \sigma(W(s_t \oplus h_t) + b) \quad (3.4)$$

Cross entropy can be used to measure the difference between the probability distribution of the real tag and the predicted result in machine learning. Therefore, it is used as the Loss function of the model, as shown in equation 3.5.

$$L = - \sum_t (a_{t+1} \log y_t^T \delta(q_{t+1}) + (1 - a_{t+1}) \log (1 - y_t^T \delta(q_{t+1}))) \quad (3.5)$$

In equation 3.5,  $a_{t+1}$  represents the true probability distribution,  $y_t^T \delta(q_{t+1})$  is the probability distribution for prediction. For the GloU\_Loss loss function, let there exist rectangles  $P$  and  $P$ , and the area where the two rectangles intersect is  $I$ , the expression of the function is 3.6.

$$GIoU = A^p + A^g - I - \frac{A^c - U}{A^c} \quad (3.6)$$

In equation 3.6,  $A^p$ ,  $A^g$  denotes the areas of rectangles  $p$  and  $g$  respectively,  $U$  denotes the area of disjoint rectangles and  $A^c$  is the area of the smallest external rectangle. To build an algorithm better suited for motor vehicle detection, the study improves the TOLOv5 network structure: the convolutional operation template in the TOLOv5 backbone network model is replaced with a Ghost template to reduce parameter redundancy during training. The related function expression is shown in Equation 3.7.

$$y_{ij} =_{i,j} (y'_i), \forall i = 1, 2, \dots, m; j = 1, 2, \dots, s \quad (3.7)$$

In equation 3.7,  $i, j$  is the linear transformation function that generates the  $j^{th}$  ghost feature map of  $y'_i$ . After reducing the amount of computational data in the model, the accuracy of the backbone network lightweight YOLOv5 model is considered to be further improved so that it can focus more on the key information to solve the task at hand amidst the large amount of input information. The study introduces Squeeze-and-Excitation (SE) and Convolutional Block Attention Module (CBAM) to improve the accuracy and efficiency of target detection. the SE schematic is shown in Figure 3.2.

As can be seen in Figure 3.2, the squeeze layer architecture of the SE is used to show the relevance of the modeled feature channels by turning the two-dimensional feature channel into a scalar with a global receptive field, and by making the shallow layer also acquire a global receptive field. In addition, the stimulus layer of SE is primarily used to generate weights for a particular channel at full connectivity to show the relevance of the modelled feature channels. The equation is shown in equation 3.8.

$$A(x) = \alpha(MLP(AvgPool(x)) + MLP(MaxPool(x))) \quad (3.8)$$

In equation 3.8  $x$  is the input feature map,  $AvgPool$  is the average pooling operation,  $MaxPool$  is the maximum pooling operation,  $MLP$  is the fully connected layer and denotes the sigmoid function. sigmoid function as a

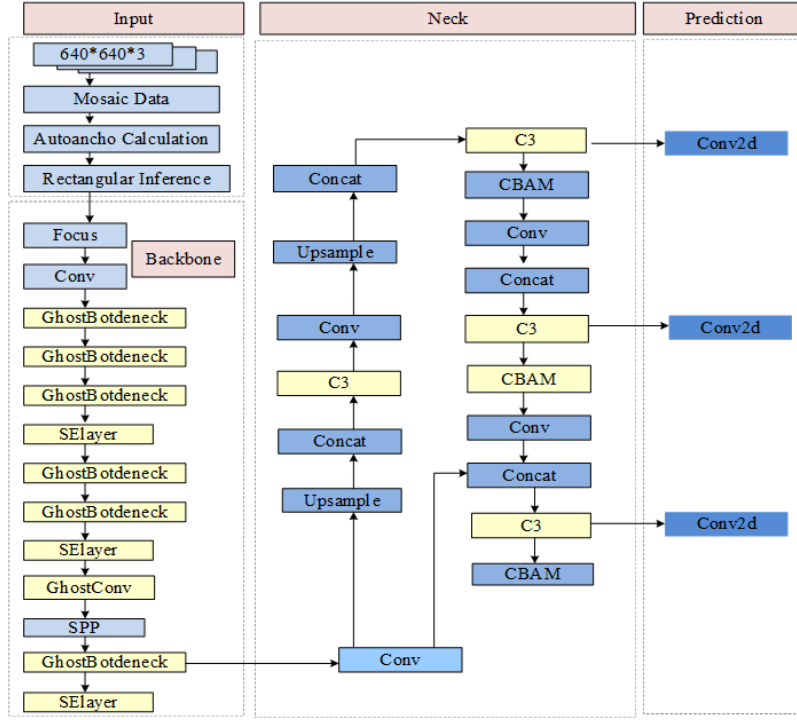


Fig. 3.3: YOLOv5-Ghost network structure diagram

weighted multiplication method with normalized weights achieves the attention threshold highlighted by the algorithm by assigning attention weights to the input features. sigmoid equation is shown in 3.9.

$$f(x) = \frac{1}{1 + e^{-x}} \quad (3.9)$$

The corresponding replacement templates are Ghost Bottleneck module and Ghost Conv module. One of the Ghost Bottleneck modules consists of two Ghost modules stacked twice. The backbone network is then given the feature map convolution operation. The CABM template is added into the lower adoptive layer of the neck to patch the deficiency of uneven weight distribution of the SE module, thus enhancing the feature representation capability. Figure 3 depicts the YOLOv5-Ghost network's organizational structure [15].

In Figure 3.3, the constructed YOLOv5-Ghost replaces the original BottleneckCSP in the neck with a C3 structure, thus making the constructed YOLOv5-Ghost model structure more efficient and concise, the network structure is also more simplified, and the resulting parameter operations involved are reduced accordingly.

**3.2. Model Construction of Multi-objective Motor Vehicle Tracking and Violation Recognition Algorithm.** Simple online Realtime tracking with a deep association metric (DeepSort) is a multi-target tracking algorithm based on target detection and is suitable for multi-target portrait tracking. DeepSort uses cascade matching to reduce the number of target ID jumps and retains the simple online Realtime tracking (SORT) algorithm for the transport characteristics of a matched target. Similarity. The covariance calculation equation is shown in Equation 3.10.

$$d_{i,j} = (d_j - y_i)^T \frac{1}{S_i} (d_j - y_i) \quad (3.10)$$

In equation 3.10,  $d_j$  is the position of the detection frame of the  $j^{th}$  target,  $y_i, S_i$  denotes the tracking position of the  $i^{th}$  Kalman filter on the target, and  $S_i$  denotes the covariance matrix. The standard deviation between

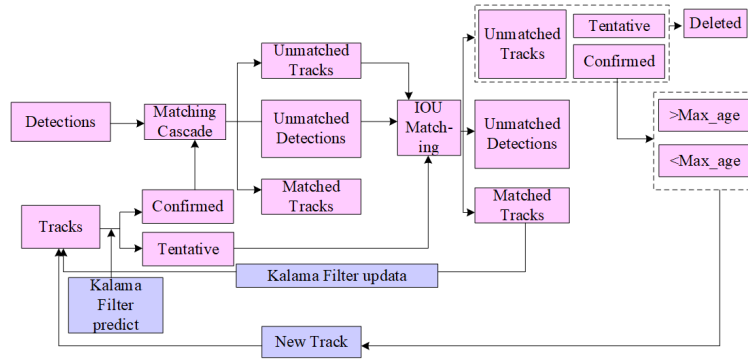


Fig. 3.4: Flowchart of Deep Sort

the mean value of the target's tracking position by the Kalman filter and the detection frame is used to estimate the uncertainty between the measured and true values of the target. Additionally, a specific covariance distance value is set to filter out targets that are not correlated. This operation will also involve the calculation of the cosine distance, and the relevant equation is shown in equation 3.11.

$$h_{(i,j)} = \min\{1 - r_j^T r_k^{(i)} | r_k^{(i)} \in R_i\} \quad (3.11)$$

In equation 3.11,  $r_j^T r_k^{(i)}$  is the cosine similarity and  $h_{i,j}$  is the minimum cosine distance between the set of vectors closest to the location tracked by the  $i^{th}$  Kalman filter and the feature vector of the  $j^{th}$  detection result. As  $d_{i,j}$  and  $h_{i,j}$  can only provide the possible positions of moving objects and cannot be applied to position prediction after the target has been obscured for a long period of time, this leads to the introduction of the correlation degree characterising the degree of combination of motion features and apparent features, which is expressed in the equation shown in equation 3.12.

$$c_{i,j} = \lambda d_{i,j} + (1 - \lambda) h_{i,j} \quad (3.12)$$

In equation 3.12,  $\lambda$  is the weighting factor. Rectification leads to the flow chart of DeepSort shown in Figure 3.4.

In Figure 3.4, the Deep Sort uses mainly recursive Kalman filtering to correlate frame-by-frame data, with a single hypothesis tracking algorithm at its core. While Deep Sort has good performance in multi-target portrait tracking processing, it is not directly applicable to multi-target tracking of motor vehicles. Therefore, some training of the Deep Sort algorithm with the dataset is needed to make it applicable to the deep appearance training model for motor vehicles. As the Veri-776 dataset contains a large amount of highly reproducible and multi-attribute motor vehicle data, and the captured images are realistic and unconstrained, and also annotated with information about different attributes, the study uses the Veri-776 dataset to train the Deep Sort algorithm for a deep appearance model applicable to motor vehicles. The constructed model is to a certain extent able to recognize multi-objective motor vehicle detection and tracking, but it cannot be better implemented for whether a motor vehicle has an illegal behavior. Open-Source Computer Vision Library (OpenCV) to identify basic traffic safety violations. The YOLOv5-DeepSort algorithm model by cascading OpenCV is shown in Figure 3.5.

The structure diagram for the motor vehicle traffic safety violation detection and recognition model is separated into three sections as shown in Figure 3.5. The main purpose of the YOLOv5-Ghost detection method is to enhance the performance of motor vehicle detection. This algorithm is primarily used to track multiple objectives while updating the position of road motor vehicles. And OpenCV is mainly used to analyse traffic scenes, realize digital modelling and the recognition of violations to ensure the effective recognition of motor vehicle safety violations.

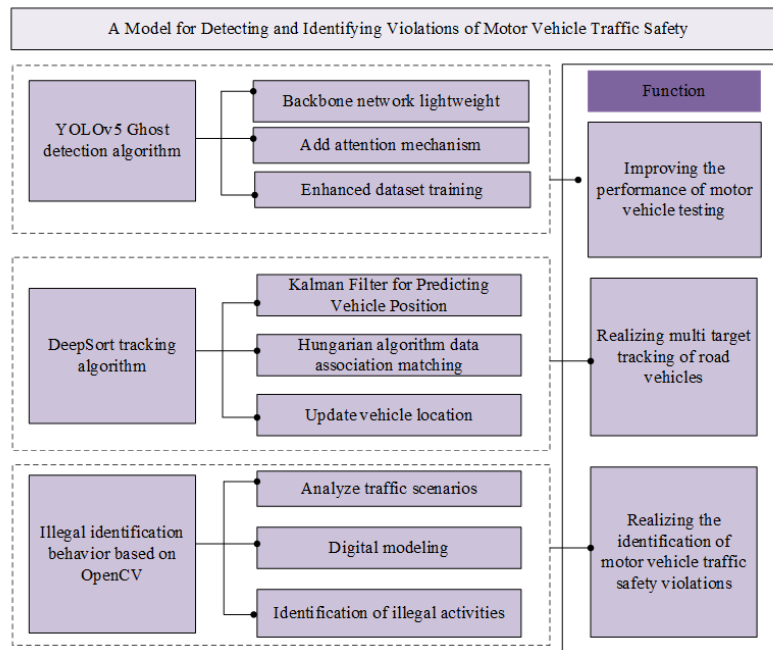


Fig. 3.5: Cascade OpenCV after YOLOv5-DeepSort algorithm model

Table 4.1: Detailed Parameters of the Server Configuration for the Experimental Run

Experimental Equipment	Experimental Parameters
System	Windows Server 2016 Standard
Memory	512GB
CPU	Intel(R) Xeon(R) CPU E7-8890v4 @ 2.20GHz
GPU	NVIDIA Quadro P6000

#### 4. Analysis of the Results of the Motor Vehicle Violation Recognition Model.

**4.1. Detection and training analysis of optimized models and other models.** The pre-processed KITTI dataset and the UA-DETRAC dataset were utilized in the study to train the motor vehicle detection algorithm in order to account for the issues with dataset format and detection target. The 89,566 total images in the two pre-processed datasets are randomly split into training and test sets in the ratio of 8:2, resulting in a total of 70,635 training images and 18,931 test images. The training environment for the YOLOv5 network was built when the dataset conversion was finished, and the specific server setup parameters for the trial run are presented in Table 4.1. To compare the experimental results of YOLOv5s, YOLOv5m, YOLOv5l, YOLOv5x, and YOLOv5-Ghost regarding accuracy, recall, weight assignment, and detection speed, tests were created to assess the performance of the YOLOv5 network structure. The dataset used for the experiments was 1565 images from the pre-processed KITTI. The results of the five network models regarding target detection are shown in Figure 4.1. As can be seen in Figure 4.1, YOLOv5Ghost has the highest accuracy and recall of 0.957 and 0.997 respectively. The number of network model layers, 91, ranks highest among the five algorithms, and the parameters required for the model calculation are 25092769. It can be said that the enhanced YOLOv5Ghost model has raised the network structure depth in comparison to the original YOLOv5, and on the basis of this, reduces the computational parameters. In terms of detection speed, the YOLOv5Ghost model has the lowest

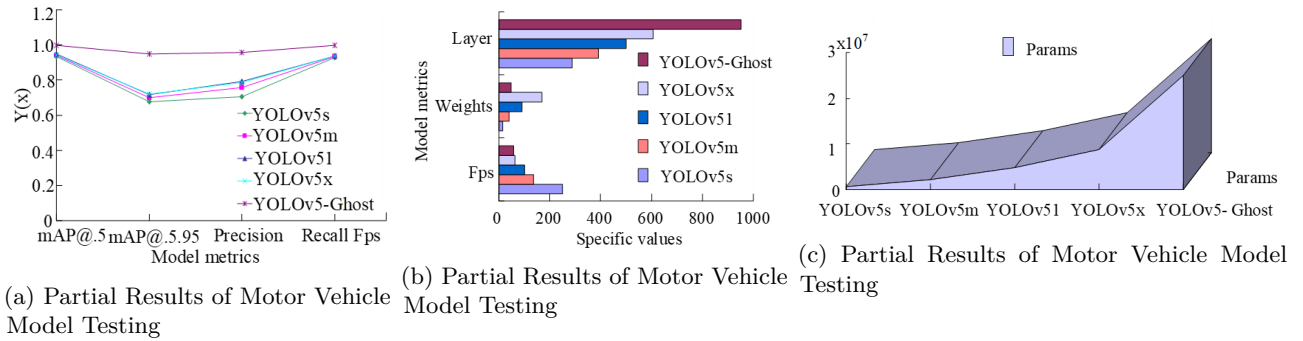


Fig. 4.1: Target detection results for five network models

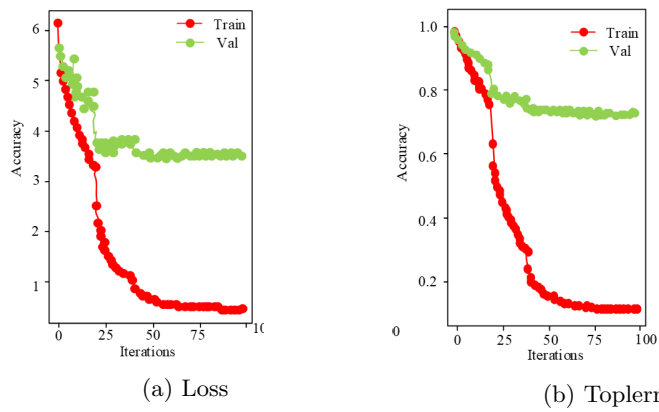


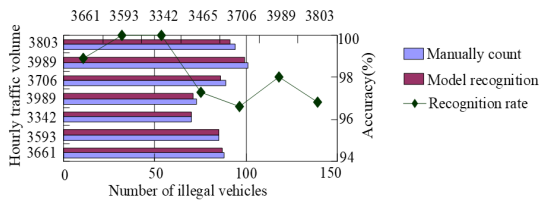
Fig. 4.2: Deep Sort training results schematic

Fps value and is not as fast as the YOLOv5 model, but has improved the accuracy and recall by up to 25.1% and 6.9%. The weight value of the improved YOLOv5Ghost model is 48.4MB, which is 42.4MB and 120MB lower than the YOLOv5l and YOLOv5x models, respectively. the reduced weight file size allows the model to run stably on embedded devices and the increased number of network detection layers reduces the operating parameters. According to the experimental findings, the Institute’s modified YOLOv5Ghost model performs better in terms of motor vehicle identification than the original YOLOv5. The Deep Sort tracking algorithm was trained using the same experimental environment as the motor vehicle tracking algorithm. The experimental outcomes are depicted in Figure 4.2. The probability and loss rate of both the training and test sets with regard to the inaccurately predicted target drop as the number of iterations rises, as shown in Figure 4.2. The decreasing trend of the LOSS curve for the training set is stable at the beginning of the iterations and increases after the number of iterations is 25. The decreasing trend of the probability of incorrect prediction curve before the number of iterations is 25 is greater than the decreasing trend after the number of iterations is 25, and the decreasing trend is again obvious when the number of iterations is 40. At around 60 iterations, the Loss curve and the probability of error prediction curve do not change significantly, indicating that the training is saturated at this point. Experiments were designed to validate the performance of the constructed YOLOv5-DeepSort motor vehicle detection and tracking algorithm.

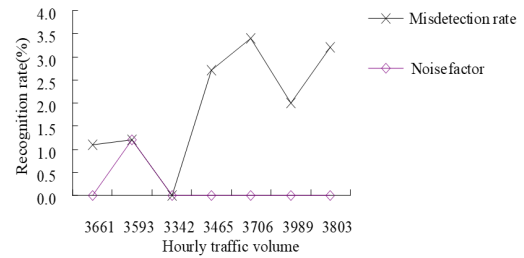


Table 4.2: The Results of Motor Vehicle Detection and Tracking Algorithms in Practical Work

Video Duration (s)	Actual No. of Vehicles	No. of Vehicles Tested	No of Missed Inspections	Misinspection Quantity
120	188	186	3	1
70	88	87	1	0
60	112	111	2	1
50	72	72	0	0



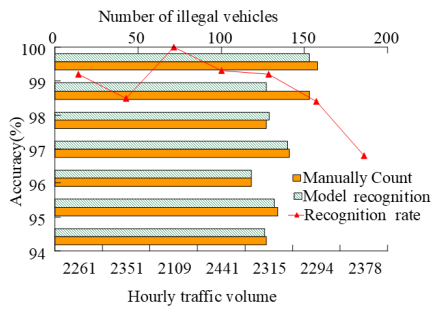
(a) Identification Results 1 of Lane Change of Motor Vehicle Compaction Line During Peak Hours of the Week



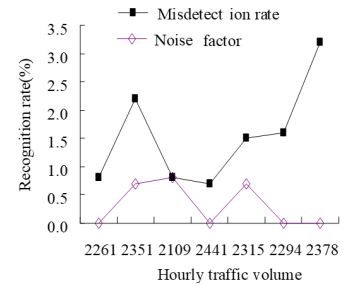
(b) Identification Results 2 of Lane Change of Motor Vehicle

Fig. 4.3: Week Peak Hour Vehicle Compaction Line Change Results

**4.2. Analysis of weekly peak recognition efficiency test for motor vehicles.** The road surface monitoring of Highway 648 and the 900m western half of the G4 Beijing-Hong Kong-Macao Expressway in Henan Province were selected as the experimental objects. The number of motor vehicles on the western half of the road section was counted manually against the number of vehicles required to be counted travelling from north to south. Table 4.2 displays the outcomes of the tracking experiment. As shown in Table 2, the Institute's proposed locomotive detection method is able to find the corresponding number of locomotives in a predetermined period of time. The best tracking experimental results are those obtained when the video duration is 50 s. The actual number of vehicles and the number of detected vehicles are the same, both being 72. The rest of the time period has a maximum error rate of 1.10% in diagnosis, corresponding to a time period of 120 s. The few cases of missed detection of vehicles that occur may be due to the fact that some of the motor vehicles have similar colors to the environment and the detection targets are small and the algorithm identifies them as pseudo-targets. False detections can arise due to interference from other dynamic objects in the surveillance vision, such as small vehicles obscuring large vehicles, thus allowing large vehicles to be identified as multiple detection targets and not being successfully merged together. After verifying the recognition and tracking effectiveness of the motor vehicle detection model, the performance of the cascaded OpenCV algorithm for determining motor vehicle traffic violations needs to be verified. The basic traffic safety violations can be categorized into three types of motor vehicle pressure realization lane change, pressure in the form of guiding lanes and intersection section vehicle turnaround violations. For the three different basic traffic safety violations, different field data are selected for the experiments. For the detection of motor vehicle lane changing violations by pressing the solid line, the real-time surveillance video of the urban section of Xiuxiang Avenue in Nanning City, Guangxi Province, from June 27, 2021 to July 3, 2021, during the period from 16:30 to 17:30 p.m. was selected for the experiment to compare the manual statistics of the weekly peak hour motor vehicle lane changing results by pressing the solid line, and Figure 4.3 displays the outcomes of the experiment. In Figure 4.3, the experimental set of statistics traffic flow can reach 3651 vehicles per hour, the total number of vehicles that change lanes by pressing the solid line is 600, and the model identifies a total of



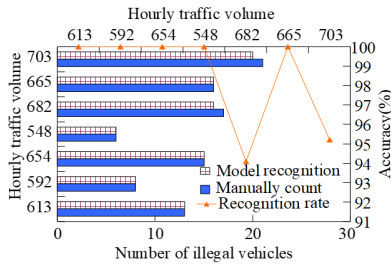
(a) Driving Results 1 of Motor Vehicle Pressure Guide Line



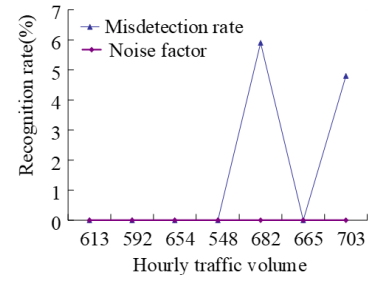
(b) Driving Results 1 of Motor Vehicle Pressure Guide Line

Fig. 4.4: Week Peak Hour Motor Vehicle Pressure Flow Line Driving Results

589 vehicles, with an average recognition rate of 98.2%. Compared to traditional manual recognition statistics, the accuracy rate of the vehicle violation recognition model constructed by the Institute ranged from 96.6% to 100%. There was no significant relationship between the recognition accuracy and peak hour traffic flow, with an average miss rate of 2.0%, a maximum miss rate of 3.4% and a minimum miss rate of 1.1%. Only one of the seven statistics included a false detection rate of 1.2%, while the false detection rate was zero in the remaining six cases. A roundabout on Chunguang Street in Fengmin District, Handan City, Hebei Province, with good weather conditions and suitable sightline conditions was selected as the subject of the experiments. The data set was selected as vehicle movement statistics from 14:30 to 15:30 in the afternoon of each day from 23 January 2022 to 29 January 2022. A UAV was used to count the relevant traffic volumes and upload the data photos to the backend. The vehicle identification tracking offence system constructed by the study was then used to compare the traditional manual statistics to produce the relevant experimental results, as shown in Figure 4.4. In figure 4.4, the average traffic flow at the roundabout was 2307 vehicles per hour. The total number of motor vehicles manually counted driving against the guideway was 937, and the model identified 925, with an average recognition accuracy of 98.7%. Again there was no significant relationship between peak hour traffic and recognition accuracy, with the highest model recognition accuracy of 100% and the lowest of 96.8% for the seven time periods. Each time period is subject to a miss detection rate, with the minimum being 0.7%. The relatively high false detection rate may be due to some vehicles driving along the edge of the guide line and the model deciding that they are pressing the line, while the high false detection rate is due to two vehicles driving along the guide line at the same time and the vehicle behind them not being effectively identified due to obscuration. The same weather conditions and good view conditions were selected, and the intersection of Bei Er Wei Yi Road in Siping City, Jilin Province was used as the experimental object. The traffic flow from 16:30 to 17:30 in the afternoon of each day from 2021.3.22 to 3.28 was selected as the data set, and experiments were designed to verify the performance of the model regarding motor vehicle illegal U-turn recognition, again comparing the traditional manual statistical results, and Figure 10 displays the outcomes of the experiments. In Figure 4.5, the traffic flow at the intersection averaged 637 vehicles/hour after manual counting, with a total of 96 vehicles in violation, and the model identified a total of 94 vehicles in violation, with an average recognition accuracy of 97.9%. Compared to the other two basic traffic violations, the model's violation recognition accuracy is higher. The recognition accuracy was 100% in five of the seven time periods. The average rate of missed detection was 2.1%, and the average rate of erroneous detection was 0%. The reason for the missed detection was that the detector was unable to precisely identify the motor vehicle in issue since some vehicles were severely veiled. We selected Chenghua Avenue in Chengdu, Sichuan Province as the experimental object, and selected daily traffic from 18:30 to 21:00 on March 11, 2023 to March 16, 2022 as the dataset. We designed an experimental validation model and provided feedback information on the vehicle system for motor vehicle violations, as shown in Figure 4.6.



(a) Recognition Results 1 of Illegal Turning of Motor Vehicles



(b) Recognition Results 1 of Illegal Turning of Motor Vehicles

Fig. 4.5: Week peak hour motor vehicle turnaround violation identification results

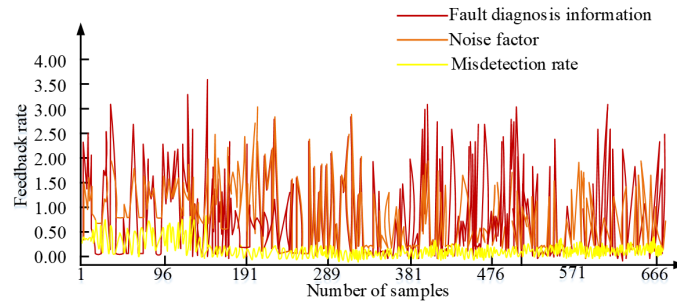


Fig. 4.6: On board system feedback information on motor vehicle violations

In Figure 4.6, when the number of samples on Chenghua Avenue exceeds 190, there is a significant decrease in the false alarm rate; When the sample size is 381 or above, there is a significant decrease in the number of noise points, but there is still rebound. When the sample size is between 230 and 360, the prediction accuracy of fault diagnosis information will significantly decrease.

**5. Conclusion.** In order to solve the problem that manual supervision means are time-consuming and laborious in dealing with motor vehicle violations and prone to oversight omissions, the research designs a motor vehicle traffic violation identification and tracking model with an improved YOLOv5Ghost model fused with OpenCV modules. The results of the experimental performance test reveal that the modified YOLOv5Ghost model has an accuracy and recall rate of 95.7% and 99.7%, respectively, with a maximum improvement of 25.1% and 6.9% when compared to the YOLOv5 model. Comparing the performance of manual statistics and the YOLOv5-DeepSort algorithm regarding motor vehicle detection and tracking, YOLOv5-DeepSort has an accurate recognition rate of 98.7%, a missed detection rate of 1.55% and a false detection rate of 0.67%. The performance of the motor vehicle recognition model incorporating the OpenCV module is verified by comparing manual statistics in a real-world environment. The experimental results show that in the experiments of changing lanes by pressing the solid line, there are 600 vehicles in total in the manual solvent and 589 vehicles in total are identified by the model, with an average recognition rate of 98.2%, an average miss detection rate of 2.0% and an average false detection rate of 0.17% by the model. The average recognition accuracy of the model was 98.7% and the average false detection rate was 0.31% in the experiment of driving on the pressure guide line. In the experiments on illegal motor vehicle U-turn identification, a total of 96 vehicles were counted manually and 94 vehicles were identified by the model, with an average recognition accuracy rate of 97.9%. The research results indicate that the modified YOLOv5Ghost model of Deepsort performs better than the original

YOLOv5 model in motor vehicle recognition. The constructed motor vehicle violation recognition model has good recognition rate and stable performance, which can save a certain amount of manpower and financial resources compared to traditional manual statistics. The YOLOv5 Ghost model structure constructed is more efficient and concise, and the network structure is also more simplified, resulting in a corresponding reduction in parameter operations.

**Funding.** The research is supported by: Science and Technology Research Project of Science and Technology Department of Henan Province, No.: 222102210089, Intelligent Decision-making System of Traffic Accident Liability Identification based on causal model and cloud service technology; Teaching Reform Project of Henan Police College, No. JY2021002, Discussion on Reality Teaching Mode of Public security Specialty.

#### REFERENCES

- [1] Nsugbe, E. Toward a Self-Supervised Architecture for Semen Quality Prediction Using Environmental and Lifestyle Factors//Artificial Intelligence and Applications. 2023. (0)
- [2] Febraro, D., Angela & Davide, N. management system for smart road networks reserved for self-driving cars. *ET Intelligent Transport Systems*. **14**, 2020 (0)
- [3] X. Yang, M. You, and Hongxing, M. *Intelligent Operational Supervision System for Operability and Reliability Analysis of Operators Manual Actions in Task Implementation. Process Safety and Environmental Protection, Part B*, **2022**, 158(0).
- [4] Sungan Yoon, J. Cho. *Onvergence Of Stereo Vision-Based Multimodal YOLOs For Faster Detection Of Potholes*. **2022** pp. 2821-2834 (0)
- [5] Changhong Pan, M. Liu,Guolong Shi.Optimization of Intelligent Logistics Supply Chain Management System Based on Wireless Sensor Network and RFID Technology. *Journal Of Sensors*. **2021**, 3-81119 (0)
- [6] Zhenxin Yao, X. Song,Lu Zhao,Yanhang Yin. *Eal-time Method For Traffic Sign Detection And Recognition Based On YOLOv*. **235**, 1978-1991 (0)
- [7] Fraser, M. & Meuleners, L. B. *Isk Factors For Unsafe Events Involving A Motor Vehicle For Group Riders (cyclists): A Naturalistic Case-crossover Study*. **146**, 3-10575 (0)
- [8] Zhang, W., Guili & Chenchen Lin. *Etecting Small Chinese Traffic Signs Via Improved YOLOv*. **2021**, 4-88265 (0)
- [9] Azimjonov, J. & Ozmen A., A. real-time vehicle detection and a novel vehicle tracking systems for estimatingand monitoring traffic flow on highways. *Dvanced Engineering Informatics*. **2021** pp. 393-402 (0)
- [10] Ahmed, A. & Misbah Gwanggil. *Op View Multiple People Tracking By Detection Using Deep SORT And YOLOv*. **12**, 3053-3067 (0)
- [11] Duan, C. & Li, X. Multi-target Tracking Based on Deep Sort in Traffic Scene. *Journal Of Physics: Conference Series*. **1952**, 22074-02208 (2021)
- [12] Huiyuan Fu, J. Guan,Feng Jing,Chuanming Wang,Huadong Ma. *Real-time Multi-vehicle Tracking Framework In Intelligent Vehicular Networks*. **18**, 89-99 (0)
- [13] Dong, Y. & Shuai, C. lightweight vehicles detection network model based on YOLOv5. *Ngineering Applications Of Artificial Intelligence: The International Journal Of Intelligent Real-Time Automation*. **11**, 2022 (0)
- [14] Wei Jia, S. Xu,Zhen Liang,Yang Zhao,Hai Min,Shujie Li,Ye Yu. *Eal-time Automatic Helmet Detection Of Motorcyclists In Urban Traffic Using Improved YOLOv*. **15**, 3623-3637 (0)
- [15] Tan, Y. & Jun Li. *Pproach For Improving YOLOv*. **15**, 4-03651 (0)

*Edited by:* Zhenling Liu

*Special Issue on:* Cloud Computing for Intelligent Traffic Management and Control

*Received:* Jun 8, 2023

*Accepted:* Aug 1, 2023



## APPLICATION OF IMPROVING ABC IN COLD CHAIN LOW CARBON LOGISTICS PATH PLANNING

**Abstract.** The market has set higher efficiency and environmental requirements for cold chain logistics, and path planning plays an important role. This study proposes a low-carbon cold chain logistics path planning model based on an improved artificial bee colony algorithm (this paragraph refers to "fusion algorithm"). The study first establishes the fusion algorithm. Then, in response to the shortcomings of this algorithm, the artificial fish swarm algorithm and genetic algorithm are used to improve it. The final results express that the shortest distance for solving Eil51 using this algorithm is 421.38, the longest distance is 448.58, and the average distance is 439.34; The shortest distance for solving Ulysses22 is 72.46, the longest distance is 73.63, and the average distance is 72.84. The average convergence times for Eil51 and Ulysses22 are 133.57 and 7.86, and the optimal performance ratios for relative error are 0.0076 and 0.0051. The robust performance ratios are 0.0362 and 0.0117. The optimal total cost solution and the average value for solving the relevant distribution problem are 47,894.6 yuan and 48,562.7 yuan, respectively. In summary, the model proposed in the study has good application effects in cold chain low-carbon logistics path planning, and has a certain promoting effect on the development of cold chain logistics.

**Key words:** ABC algorithm; Cold chain low-carbon logistics; Artificial fish school algorithm; Path optimization; Genetic algorithm

**1. Introduction.** As the market grows and consumer demands continue to evolve, the cold chain logistics (CCL) industry faces more opportunities and challenges. Distribution centers need to minimize food consumption, reduce costs and carbon emissions while meeting consumer needs, and path planning plays a critical role in this process. Cold Chain Low-Carbon Logistics Path Planning (CCLCLPP) is a variant of the Vehicle Routing Problem (VRP), which refers to planning suitable paths for a set of loading and unloading points. Under certain constraints, the achievement of certain objectives and the use of efficient routes provide transportation companies with a direct competitive advantage. The Artificial Bee Colony Algorithm (ABC) is inspired by the process of bees picking honey. Due to its fast computing speed and high accuracy, it has occupied a vital position in solving VRP in recent years. However, ABC also has the drawbacks of slow convergence speed and a tendency to fall into local optima. Some scholars have used SURF, RANSAC, and particle swarm optimization algorithms to improve the ABC algorithm [3, 12]. To improve the transportation efficiency and economic benefits of cold chain logistics, reducing transportation time plays a crucial role. Proper planning of the transportation route of cold chain logistics can greatly reduce the transportation time. In this context, the Artificial Fish School (AFS) and Genetic Algorithm (GA) were utilized to improve the ABC algorithm and construct a CCLCLPP model based on AFS-GA-ABC. There are two main innovative points in this study. The first point is to introduce the clustering behavior of the AFS algorithm, while adopting a perception range based on under damping motion adaptation to compensate for the lacks of ABC, which is prone to falling into local optima and slow convergence speed. The second point is to introduce the partial mapping crossover operator in the GA to lift the slow convergence speed of the ABC. The main structure of the study is segmented into four parts: Part 1 analyzes the current relevant research status; the second part is to perfect the shortcomings of ABC by combining AFS and GA, and construct a CCLCLPP model based on improved ABC; the third part is to analyze the application effectiveness of the proposed CCLCLPP model based on AFS-GA-ABC; the final part is a summary of the entire study.

**2. Related works.** ABC is an optimization method proposed to mimic the foraging behavior of bee populations, with fast convergence speed. It is widely used to solve various optimization problems and multi-dimensional problems. The probability method used by ABC and most of its modifications selects good food sources during the foraging phase of the bee population, but its possibility selection does not take effect with increasing iterations. Ozbay E proposed an active deep learning approach using a new multi-layer structure

to automatically detect the stages of retinopathy, and used the ABC algorithm for image segmentation. The results show that the proposed method has high performance in detecting retinopathy and determining its severity from different fundus images [5]. Guo H et al. proposed a mathematical model based on the ABC algorithm to maximize the expected recovery profit in the case of task failure in real life, where scrapped products may experience different degrees of wear and tear during the disassembly process. The results show that the solution performance of this method is superior to the other three existing methods [5]. Toktas A et al. proposed for the first time a parameter chaos mapping optimization method based on three-objective optimization, and then used Pareto based three-objective ABC algorithm for optimization. The results show that the algorithm exhibits applicability in image encryption, as well as optimal traversal and complexity [13]. Sefati S et al. focused on load balancing and routing issues in wireless sensor networks, using Markov models and ABC algorithms to find the best candidate nodes for each cluster. The results show that this method outperforms the comparison method in terms of energy efficiency and the number of active nodes [9]. Yolcu V et al. optimized the wavelength and power values of pump lasers used in distributed fiber Raman amplifiers by using an adaptive ABC algorithm based on binary search equations to find the optimal pump wavelength and power level [17]. Satoh T et al. utilized the ABC algorithm to solve the design problem of a discrete-time stable unknown input estimator based on parameter optimization, and compared the proposed design method with previous design methods [8]. Wang H et al. proposed an improved multi-objective ABC algorithm based on decomposition and dimension learning to help ABC solve multi-objective optimization problems. The results indicate that the proposed method has achieved good performance [14].

CCLCLPP is a VRP developed on the foundation of CCL and environmental protection, which is quit vital for developing the CCL. Tian G et al. systematically reviewed several papers and constructed the overall structure of multi-criteria decision-making technology, in response to the lack of comprehensive review of multi-criteria decision-making in the fields of low-carbon transportation and green logistics. They also proposed the future development direction of multi-criteria decision-making technology for green logistics and low-carbon transportation systems [11]. Wang Z et al. established a multi-objective hazardous material transportation route planning model considering road traffic elasticity and low-carbon to address the increasing proportion of hazardous materials in domestic road transportation, filling the gap in research on hazardous material transportation in the low-carbon field [15]. Tao N et al. established a mathematical model for optimizing the path of cold chain logistics delivery vehicles with the lowest comprehensive cost. The proposed improved hybrid ant colony optimization algorithm solves the problems of increasing difficulty in path optimization and carbon emissions in the cold chain logistics distribution process. [10]. Raman P et al. studied low-carbon performance based on low-carbon supply chain practices in the manufacturing industry. The research results indicate that low-carbon production is insignificant in reducing overall carbon emissions and can be used to develop and extend a low-carbon supply chain framework [7]. Arora M et al believe that due to the various challenges faced by the entire frozen food industry, it is necessary to study in India. So they reviewed the existing literature in an attempt to explain the issues that affect frozen food in India and strategies to address these issues. They also proposed a conceptual model and described the relationship [1]. Xu L et al. systematically reviewed the current academic literature on the role of technology in low-carbon supply chain management and provided a novel and comprehensive roadmap for future research on technology-enabled low-carbon supply chains. [16]. Cheng C et al. proposed a selection standard from the perspective of low-carbon level to address the issue of how enterprises can select a logistics supplier that can provide low-carbon and high-quality services [4].

In summary, although many previous scholars and scientists have recognized the importance of CCLCLPP in reducing logistics costs, improving customer satisfaction, and protecting the environment. ABC has played a significant role in solving various optimization problems and has been proven, but the effectiveness of the ABC-based CCLCLPP is still not ideal. To make up for this deficiency, studying the AFS and GA algorithms to improve the ABC algorithm has important practical application value and prospects for CCL.

**3. Construction of CCLCLPP model based on improved ABC.** The ABC can be taken to deal with various optimization matters and has certain effectiveness in the application of traditional CCL path planning. But with the improvement of users' demands, the market has put forward higher efficiency, economy, and low-carbon requirements for cold chain transportation. The path planning method for CCL should also be updated with the times. Therefore, a ABC-based CCLCLPP algorithm was studied and constructed, and

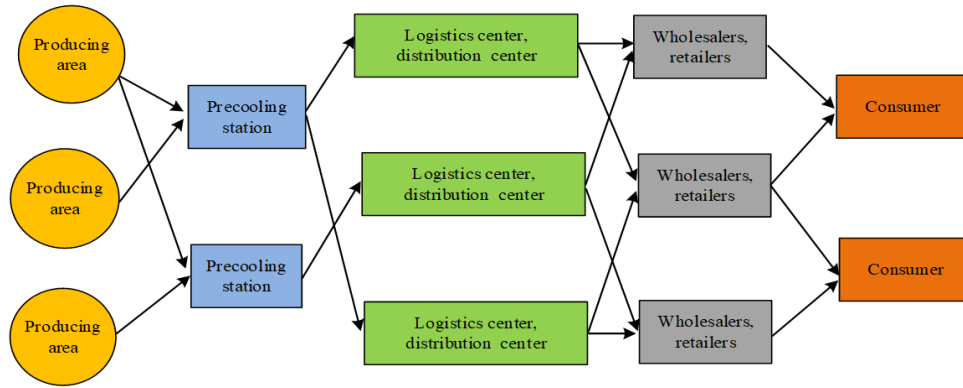


Fig. 3.1: Cold chain logistics network structure

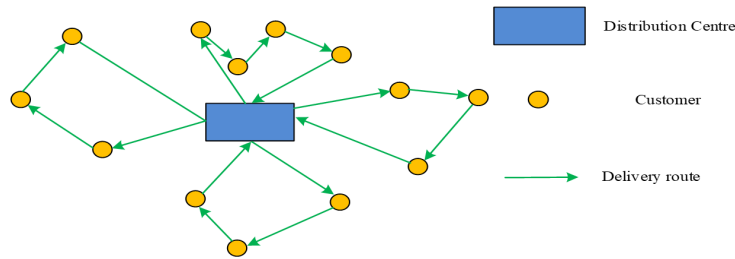


Fig. 3.2: Schematic diagram of vehicle routing issues

the traditional ABC algorithm was improved using AFS and GA to build a new CCLCLPP model grounded on improved ABC.

**3.1. Construction of CCLCLPP model based on ABC.** According to the definition of the European Union, CCL means a systematic engineering process in which refrigerated foods are always kept in a specified low-temperature space from production to consumption, to reduce food loss and ensure food quality. It involves multiple technological fields such as information technology, modern refrigeration technology, logistics technology, and has the characteristics of high cost, high timeliness, high carbon emissions, easy product wear and tear, and complex technology [2]. The structure of CCL is listed in Figure 3.1.

The CCLCLPP problem is an evolution of the VRP problem. To reduce the delivery cost of refrigerated and frozen food, improve the delivery efficiency, reduce the carbon emission, and meet the basic needs of consumers, it is necessary to scientifically plan the route of delivery vehicles. The Figure 3.2 shows the diagram of VRP.

The VRP problem includes five basic characteristic elements: delivery vehicles, delivery centers, consumers, delivery routes, and optimization objectives. Its classical mathematical model is Equation 3.1.

$$MinZ = \sum_{i=0}^n \sum_{j=1}^n \sum_{k=1}^K cd_{ij}x_{ij}^k \tag{3.1}$$

In Equation 3.1,  $K$  is the amount of vehicles,  $n$  is the consumer numbers,  $c$  represents the transportation cost per kilometer, and  $d_{ij}$  means the distance between  $i$  and  $j$ . When  $x_{ij}^k$  is equal to 1, it means that vehicle  $k$  arrives at consumer  $j$  from consumer  $i$ , and in other cases,  $x_{ij}^k$  is equal to 0. ABC algorithm was originally used to solve function optimization problems, and now it has been applied to data processing, image processing, cold

chain logistics, and other fields. However, it is easy to fall into local optimization and has a slow convergence rate. Therefore, ABC algorithm is selected as the basic algorithm for this study and further optimized. The ABC is superior in fewer parameters, convenient calculation, easy implementation, and strong robustness. It is widely used to solve multidimensional problems and model optimization problems, as well as to solve path planning problems. ABC divides the foraging bees into 3 types: leading, following and scouting bees, and sets the maximum cycle times and the limit value of the times that the leading bees do not update the food source to convert to the scouting bees. The initialization bee colony is a honey source of randomly generated  $n$ -dimensional vectors, and the generation of each honey source is Equation 3.2.

$$\begin{cases} x_{ij} = x_{\min j} + \text{rand}(x_{\max j} - x_{\min j}) \\ i = 1, 2, \dots, m, j = 1, 2, \dots, n \end{cases} \quad (3.2)$$

In Equation 3.2,  $x_i$  is the  $i$ -th honey source in the bee colony,  $\text{rand}$  represents a random number uniformly distributed between  $(0,1)$ .  $x_{\min j}$  and  $x_{\max j}$  represent the lower and upper bounds of vector  $j$ . After initialization, calculate the fitness value of the honey source, see Equation 3.3.

$$\text{fit}(x_i) = \begin{cases} 1 + f(x_i), f(x_i) \geq 0 \\ \frac{1}{1+|f(x_i)|}, f(x_i) \leq 0 \end{cases} \quad (3.3)$$

In Equation 3.3,  $f(x_i)$  expresses the concentration of the  $i$ -th honey source. In the leading bee stage, each leading bee conducts a domain search around the current honey source to find a new one, see Equation 3.4.

$$v_{ij} = x_{ij} + \phi_{ij}(x_{ij} - x_{kj}), \quad k = 1, 2, \dots, SN \quad (3.4)$$

In Equation 3.4,  $v$  represents the new honey source,  $k$  and  $j$  represent the randomly selected index, and  $\phi_{ij}$  represents the random number between  $[-1,1]$ . When a new honey source is found, the leading bee evaluates it based on the principle of greedy selection and compares it with the old honey source. If the new honey source has a higher fitness, the old one is replaced by the new one, otherwise it remains unchanged. Later, the following bee evaluates all the information received from the leading bee and chooses a honey source, if possible, depending on the fitness value of the honey source in the population. The selection mechanism based on fitness can adopt the roulette wheel algorithm, and the calculation method is Equation 3.5.

$$p_i = \frac{f_i}{\sum_{j=1}^m f_j} \quad (3.5)$$

In Equation 3.5,  $f_i$  represents the fitness value of the  $i$  honey source, and the greater the  $f_i$ , the greater the probability of honey source being selected. A CCLCLPP model was designed using the ABC algorithm, and the solution process of this model is Figure 3.3.

However, ABC has drawbacks in solving complex problems, such as too many iterations, tendency to fall into local extreme traps, low optimization accuracy, and slow speed. To ensure the actual operational effect of the CCLCLPP model, further optimization of the ABC algorithm is needed.

**3.2. Construction of CCLCLPP Model Based on AFS-GA-ABC.** With the growing needs of the cold chain transportation industry and the implementation of green concepts, improving the CCLCLPP method is becoming increasingly important. To compensate for the shortcomings of the ABC algorithm, which is prone to local optima and slow convergence speed, the clustering behavior of the AFS algorithm is introduced, and a perceptual domain based on underdamping motion adaptation is adopted. AFS algorithm is a bionic algorithm that realizes optimal target search by simulating fish feeding, tail chasing, clustering and other methods, and realizes global optimization by local optimization. It is flexible with fast conversion and insensitive to the original parameter settings [6]. The specific process is shown in Figure 3.4.

When implementing the AFS algorithm, it is necessary to assign a perception area to each bee and to determine the perception distance between bees using the following Equation 3.6.

$$\begin{cases} d_i = d_i + 1 \\ X_{ik} \neq X_{jk} \end{cases}, \quad k = 1, 2, \dots, N, \quad i \neq j \quad (3.6)$$



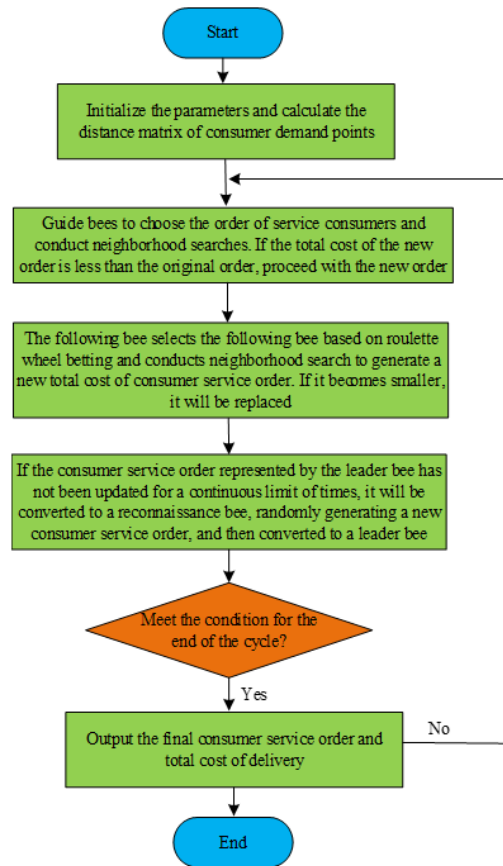


Fig. 3.3: Solution Process for the Optimization Model of ABC-based CCLCLPP

In Equation 3.6,  $X_{ik}$  and  $X_k$  are the path order of the  $i$ -th and  $j$  bee.  $k$  represents the vector dimension.  $N$  is the amount of leading peaks. The calculation of the bee amounts with a perception distance less than the perception range is Equation 3.7.

$$\begin{cases} p_i = p_i + 1 \\ di \leq Visual \end{cases} \quad (3.7)$$

In Equation 3.7,  $Visual$  represents the perceptual range of bees. Then determine whether the bees with a perception distance less than the perception range are crowded around them, as shown in Equation 3.8.

$$\frac{p_i}{N} \leq \det al \quad (3.8)$$

In Equation 3.8, the quantity of leading, following, and reconnaissance bees is the same, and is the crowding factor. In the leading bee stage, when the center position is not crowded, a clustering behavior with the characteristic of accelerating convergence speed is introduced for neighborhood search. At the same time, the calculation method of the bee's perception distance is introduced into the underdamped motion in physics by using an adaptive update strategy. In the early stages of iteration, it is better to have a larger perception distance; in the later stages of iteration, it is better to have a smaller perception distance. The motion law of the spring oscillator in underdamped motion under the condition of no constraining force is Equation 3.9.

$$\begin{cases} X = A^0 \cdot e^{\delta t} \cos(\omega t + \phi) \\ 2\delta = \sqrt{\frac{g}{m}} \end{cases} \quad (3.9)$$

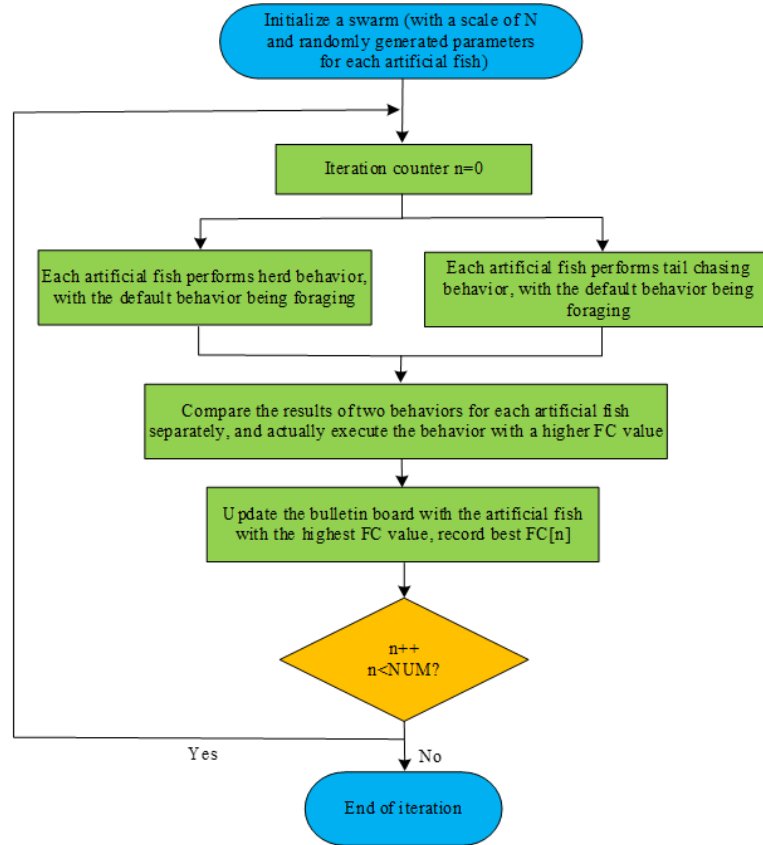


Fig. 3.4: Diagram of AFS

In Equation 3.9,  $\omega$  represents the damping factor,  $t$  is time,  $\phi$  represents vibration frequency, and  $m$  is the mass of the oscillator. The adaptive update strategy for designing the perception range based on the damping motion law is Equation 3.10.

$$Visual = 2 + \frac{10 \log(ite\text{r})}{\log(Maxite\text{r}ation)} \times \cos\left(0.5\pi - 0.5\pi \frac{Maxite\text{r}ation - ite\text{r}}{Maxite\text{r}ation}\right) \quad (3.10)$$

In Equation 3.10,  $Maxite\text{r}ation$  is the max iterations, and  $ite\text{r}$  is the current iterations. The GA algorithm is an optimization algorithm that simulates phenomena such as crossover, replication, and mutation during the evolution of species, and is often used to solve optimization problems. To improve the slow convergence speed of ABC, some mapping crossover operators are introduced into the GA algorithm for optimization. In the leading bee stage, if the center position is crowded, a new solution is obtained by the crossover operation based on the partial mapping crossover operator to reduce the probability. Later, if the neighborhood search results in a new solution that is not excellent than that of the original leading bee, this new solution is obtained by performing a crossover operation based on the partial mapping crossover operator. The partial mapping crossover operator is mainly divided into two steps: partial crossover operation and generation of new individuals based on the corresponding mapping relationship. The specific process is shown in Figure 3.5.

The solution of the CCLCLPP model based on AFS-GA-ABC is divided into 5 steps. First, the parameters of the improved ABC are initialized and the distance matrix of consumer demand points is calculated. Second, in the leading bee stage, the leading bee first flies out of the hive to select the order of serving consumers, and conducts neighborhood search to compare the total cost of the new order with the original order. If the cost

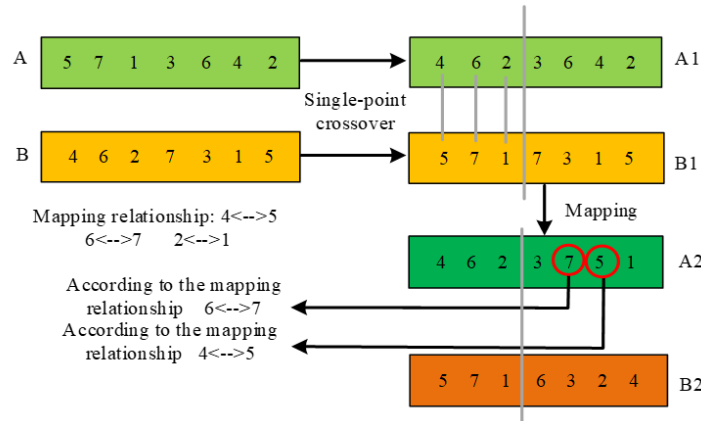


Fig. 3.5: Partial mapping crossover operator flowchart

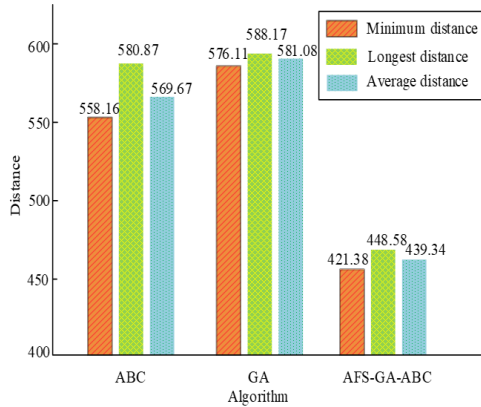
decreases, a new order is performed, otherwise clustering behavior is performed. If the clustering behavior is successful and the total cost decreases, it is replaced, otherwise it remains unchanged. If the clustering behavior fails, a new service order is obtained by performing a crossover operation based on the partial mapping crossover operator, which is the same as above. Third, in the follower bee stage, the follower bee selects the lead bee to follow based on the roulette wheel bet, and performs neighborhood search to generate a new total cost of consumer service order. If it becomes smaller, replace it; otherwise, perform the crossover operation based on the partial mapping crossover operator to obtain a new order; If the total cost decreases, replace it, otherwise it remains unchanged. Fourth, if the consumer service order represented by a leader bee has not been continuously updated for a limit number of times, the leader bee is transformed into a scout bee. Upon entering the scout bee stage, the scout bee randomly generates new consumer service orders, and then the scout bee turns into a leader bee. Fifthly, determine whether the conditions for the end of the cycle have been met. If not, repeat steps two through four of the cycle until the end-of-cycle conditions are met, and output the end user service order and total delivery cost.

**4. Effect Analysis of AFS-GA-ABC Cold Chain Low Carbon Logistics Path Planning Model.**

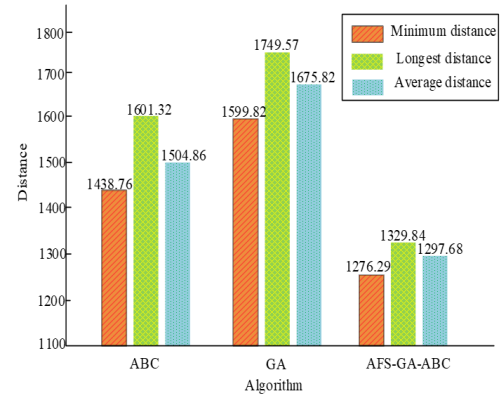
The AFS-GA-ABCCCLCLPP model is used to plan reasonable distribution routes, which is beneficial for improving distribution efficiency and reducing distribution costs. This has a certain positive significance for the promotion of CCL, but the actual application effect of this model still needs further verification. The research mainly analyzes from two aspects. The first part is to conduct simulation experiments and analysis on the AFS-GA-ABC algorithm, and the second part is to analyze the practicability effect of it in the CCLCLPP.

**4.1. Simulation Experiment Analysis of AFS-GA-ABC.** To verify the effectiveness of the AFS-GA-ABC, Eil51 and Ulysses22 were selected for simulation experiments. The experiment verifies the performance of the algorithm using four indicators: path, convergence times, optimal performance ratios, and robust performance ratios. The shorter the path, the better the optimization effect of path planning. The shorter the convergence time, the higher the efficiency. The lower the optimal performance ratios and the lower the robust performance ratios, the better the stability of the algorithm. And compares it with the ABC and GA. The amount of leading bees and following bees are set to 50; is 70; is 500; The max attempt number in group behavior is 250; The probability of partial mapping crossover is 0.8; The crowding coefficient is 0.625. The shortest known distance for Eil51 is 419, and the experiment was run 25 times separately. To use three algorithms to solve Eil51 and compare them with Rat99, which solves for the shortest distance of 1205 4.1. The shortest distance for AFS-GA-ABC to solve Eil51 is 421.38, the longest distance is 448.58, and the average distance is 439.34, both lower than the results obtained by ABC and GA. The best algorithm to solve Eil51 and Rat99 is AFS-GA-ABC.

The performance of the three algorithms for solving Eil51 is shown in Table 4.1. The average convergence



(a) The Results of Solving Eil51



(b) The Results of Solving Rat99

Fig. 4.1: The results of solving Eil51 and Rat99 using three algorithms

Table 4.1: Eil51 Solves the Optimal Performance Ratio

Algorithm	Avg. Convergence Times	Relative Error Optimal Perf. Ratio	Robust Perf. Ratio
ABC	428.37	0.3186	0.3396
GA	365.87	0.3576	0.3654
AFS-GA-ABC	133.57	0.0076	0.0362

number obtained by AFS-GA-ABC for Eil51 is 133.57, with the best relative error and robustness performance ratios of 0.0076 and 0.0362, both of which are better than the results obtained by ABC and GA solutions.

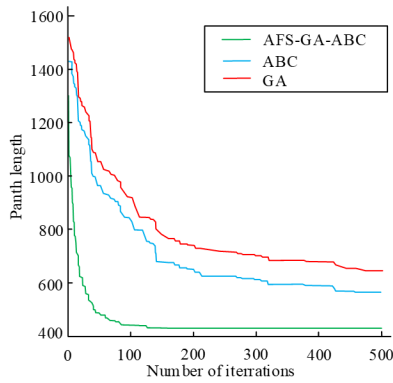
The performance of the three algorithms in solving Eil51 is shown in Table 4.1. The average convergence number obtained by AFS-GA-ABC for Eil51 is 133.57, with the best relative error and robustness performance ratios of 0.0076 and 0.0362, both of which are better than the results obtained by ABC and GA solutions.

The shortest known distance is 72, and the experiment is run 25 times separately. Three algorithms were used to solve Ulysses22 and compared to KroB100 with the shortest distance of 22136 as shown in Figure 4.3. The shortest distance obtained by solving Ulysses22 using AFS-GA-ABC is 72.46, the longest distance is 73.63, and the average distance is 72.84, both lower than the results of ABC and GA; Moreover, the most effective algorithms for solving Ulysses22 and KroB100 are AFS-GA-ABC.

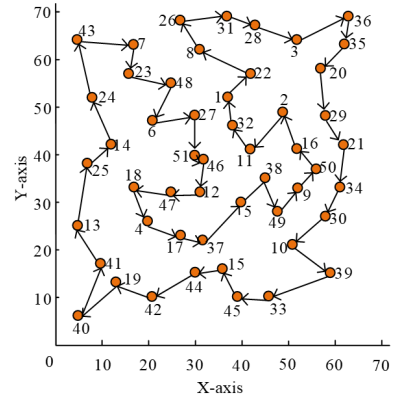
Table 4.2 shows the comparative results of three algorithms for solving Ulysses22. The average convergence number obtained by using AFS-GA-ABC to solve Ulysses22 is 7.86, the optimal relative error performance ratio is 0.0051, and the robust performance rate is 0.0117, both of which are better than the results of ABC and GA.

Figure 4.4 shows the optimization curves and shortest paths of three algorithms for Ulysses22. The path length obtained by AFS-GA-ABC for Ulysses22 and the iterations required to obtain it are much smaller than those using ABC and GA, therefore AFS-GA-ABC has higher accuracy and efficiency.

In summary, AFS-GA-ABC solves a shorter shortest path compared to ABC and GA, resulting in better search accuracy, higher efficiency, faster convergence speed, and better stability and robustness.

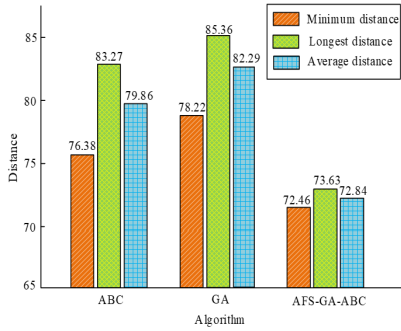


(a) Eil51 Optimal Curve

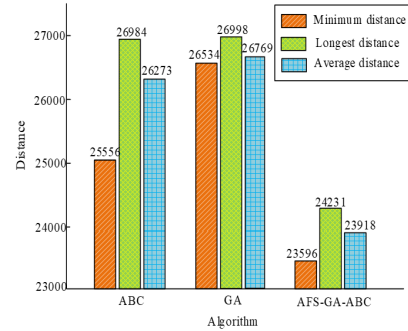


(b) Eil51 shortest path graph

Fig. 4.2: Two Algorithms for Solving Eil51's Optimization Curve and Shortest Path



(a) The results of solving Ulysses22



(b) The results of solving KroB100

Fig. 4.3: The results of solving Ulysses22 and KroB100 using three algorithms

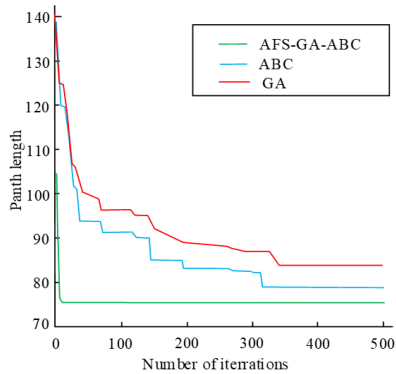
**4.2. Analysis of the Application Effect of AFS-ABC Cold Chain Low Carbon Logistics Path Planning Model.** Initializing the improved ABC parameters, setting the quantity of leading bees and following bees to 20; is 20; is 200; The max number of attempts in group behavior is 100; The probability of partial mapping crossover is 0.9; The crowding factor is 0.618. Taking one of the distribution centers of a CCL distribution company as an example, 15 consumers were provided with distribution services, and the CCLCLPP model was simulated 25 times using Simulated Annealing(SA), ABC and AFS-GA-ABC, respectively. The final total cost solution is Table 4.3. The optimum solution and mean value of the total cost gained by ABC are 58,152.4 yuan and 61,028.3 yuan, respectively; The optimum solution and mean value of the total cost gained by SA are 51213.6 yuan and 53524.3 yuan, respectively; The solutions of AFS-GA-ABC are 47,894.6 yuan and 48,562.7 yuan, both lower than SA and ABC, and the average iterations is also lower than SA and ABC. The data shows that the total cost, stability, and convergence speed obtained by using AFS-GA-ABC solution are better than SA and ABC.

In this example, the first column has a fixed width of 4cm, and the other columns have fixed widths of 2cm each. Adjust the widths as needed for your layout.

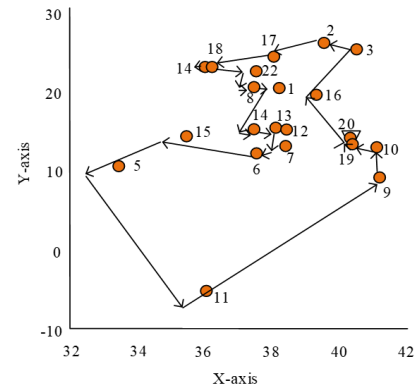
The distribution path obtained by solving the CCLCLDPP issue using ABC and AFS-GA-ABC is Figure 4.3.

Table 4.2: Ulysses22 Solves the Optimal Performance Ratio

Algorithm	Avg. Convergence Times	Relative Error Optimal Perf. Ratio	Robust Perf. Ratio
ABC	308.43	0.0597	0.1008
GA	358.26	0.0858	0.1128
AFS-GA-ABC	7.86	0.0051	0.0117



(a) Ulysses22 Optimal Curve



(b) Ulysses22 Shortest Path Graph

Fig. 4.4: Two Algorithms for Solving Ulysses22's Optimization Curve and Shortest Path

Due to the fact that AFS-GA-ABC reduces unnecessary intersections of delivery paths when solving related planning problems, the total distance of delivery paths is shorter, resulting in lower total cost.

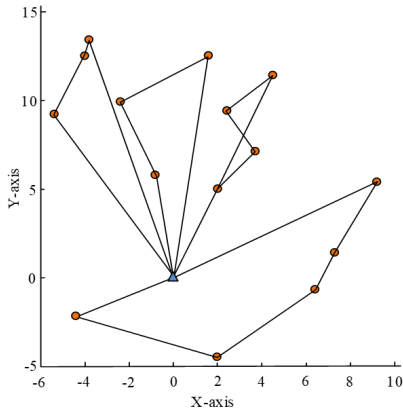
The convergence process of the two algorithms is shown in Figure 3.1. Comparison of the convergence process diagram of ABC and AFS-GA-ABC: Compared with ABC, AFS-GA-ABC has faster convergence speed and better solution effect in solving the CCLCLDPP problem.

In summary, using AFS-GA-ABC to solve path planning problems results in lower total cost, higher efficiency, and better stability compared to using ABC.

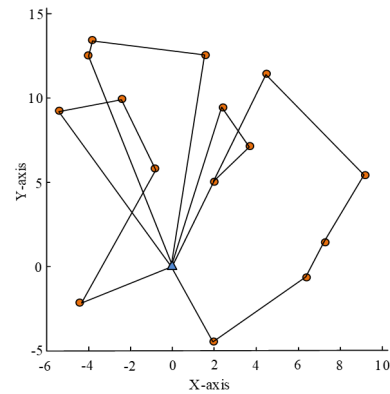
**5. Conclusion.** With the development of the market and the increasing demand of consumers, the CCL industry is facing new challenges. High efficiency, low carbon, low cost and low loss have become the mainstream development trend of CCL distribution. A CCLCLDPP model based on AFS-GA-ABC is proposed for the issue of CCLCLDPP. The conclusion is that the shortest distance for AFS-GA-ABC to solve Eil51 is 421.38, the longest distance is 448.58, and the average distance is 439.34. The shortest distance for solving Ulysses22 is 72.46, the longest distance is 73.63, and the average distance is 72.84, both of which are lower than the results of ABC. The average convergence times obtained by solving Eil51 and Ulysses22 are 133.57 and 7.86, respectively. The optimal performance ratios for relative error are 0.0076 and 0.0051, and the robust performance ratios are 0.0362 and 0.0117, both of which are better than ABC. The optimal total cost solution and the average value obtained by solving the CCLCLDPP problem are 47894.6 yuan and 48562.7 yuan, respectively, which are lower than ABC and reduce unnecessary crossing of distribution paths. The total distance of the distribution paths is shorter, and the total cost is lower. In summary, AFS-GA-ABC solves a shorter shortest path compared to ABC, resulting in better search accuracy and faster convergence speed. AFS-GA-ABC has better stability and robustness, with lower total cost and higher efficiency. However, this study only considered the problem of distribution path planning without analyzing other factors that affect distribution efficiency, cost, and carbon

Table 4.3: SA, ABC, and AFS-ABC Solution Results

Algorithm	Total Cost Optimal Solution (yuan)	Optimal Number of Iterations	Avg. Total Cost (yuan)	Avg. Convergence Algebra
ABC	58152.4	113	61028.3	169
SA	51213.6	87	53524.3	92
AFS-GA-ABC	47894.6	12	48562.7	43



(a) ABC's Delivery Path Diagram



(b) AFS-GA-ABC's Delivery Path Diagram

Fig. 4.5: Delivery Path Map

emissions. Therefore, it is necessary to consider more influencing factors, such as the weight and quantity of transported goods, vehicle model, etc., and adopt the control variables to conduct simulation experiments to better support the development of cold chain logistics.

**Funding.** The research is supported by Gansu Provincial Department of Education: University Teachers' Innovation Fund Project: Digital Economy Empowering High Quality Development of Fresh Agricultural Products Cold Chain Logistics - A Case Study of Gansu Province (No. 2023A-215).

## REFERENCES

- [1] M. ARORA, R. KUMAR, AND T. B. RAJU, *Identification of issues in the cold chain of Indian frozen food*, International Journal of Logistics Economics and Globalisation, 10 (2023), pp. 1–22.
- [2] ———, *Identification of issues in the cold chain of Indian frozen food*, International Journal of Logistics Economics and Globalisation, 10 (2023), pp. 1–22.
- [3] D. AWASTHI AND S. V. K. ROBUST, *imperceptible and optimized watermarking of dicom image using schur decomposition, lwt-dct-svd and its authentication using surf*, Multimedia Tools and Applications, 82 (2023), pp. 16555–16589.
- [4] C. CHENG, X. WANG, AND X. REN, *Selection of outsourcing logistics providers in the context of low-carbon strategies*, Environmental Science and Pollution Research, 30 (2023), pp. 18701–18717.
- [5] H. GUO, L. ZHANG, Y. REN, Y. LI, Z. ZHOU, AND J. WU, *Optimizing a stochastic disassembly line balancing problem with task failure via a hybrid variable neighborhood descent-artificial bee colony algorithm*, International Journal of Production Research, 61 (2023), pp. 2307–2321.
- [6] F. POURPANAH, R. WANG, C. P. LIM, ET AL., *A review of artificial fish swarm algorithms: Recent advances and applications*, Artificial Intelligence Review, 56 (2023), pp. 1867–1903.
- [7] P. RAMAN, F. MAHMUD, AND S. S. A. R. O. M. C. E. F. S. S. C. MANAGEMENT., *International journal of industrial management*, 17 (2023), pp. 108–118.

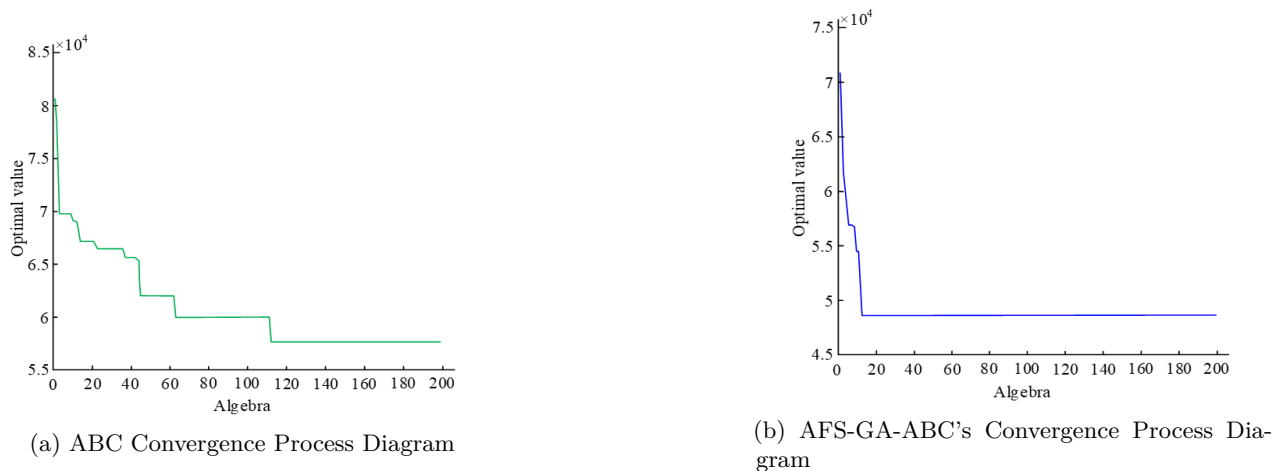


Fig. 4.6: Convergence process diagram

- [8] T. SATOH, S. NISHIZAWA, J. NAGASE, N. SAITO, AND N. SAGA, *Artificial bee colony algorithm-based design of discrete-time stable unknown input estimator*, *Information Sciences*, 634 (2023), pp. 621–649.
- [9] S. S. SEFATI, M. ABDI, AND A. GHAFARI, *Qos-based routing protocol and load balancing in wireless sensor networks using the markov model and the artificial bee colony algorithm*, *Peer-to-Peer Networking and Applications*, 16 (2023), pp. 1499–1512.
- [10] N. TAO, H. YUMENG, AND F. MENG, *Research on cold chain logistics optimization model considering low-carbon emissions*, *International Journal of Low-Carbon Technologies*, 18 (2023), pp. 354–366.
- [11] G. TIAN, W. LU, X. ZHANG, M. A. DULEBENETS, A. ALEKSANDROV, A. M. FATHOLLAHI-FARD, AND M. A. IVANOV, *survey of multi-criteria decision-making techniques for green logistics and low-carbon transportation systems*, *Environmental Science and Pollution Research*, 30 (2023), pp. 57279–57301.
- [12] S. TIWARI AND K. A. A. AND, *and bibliographic analysis of particle swarm optimization applications in electrical power system: Concepts and variants*, *Evolutionary Intelligence*, 16 (2023), pp. 23–47.
- [13] A. TOKTAS, U. ERKAN, D. USTUN, AND X. WANG, *Parameter optimization of chaotic system using pareto-based triple objective artificial bee colony algorithm*, *Neural Computing and Applications*, 35 (2023), pp. 13207–13223.
- [14] H. WANG, S. WANG, Z. WEI, T. ZENG, AND T. YE, *An improved many-objective artificial bee colony algorithm for cascade reservoir operation*, *Neural Computing and Applications*, 35 (2023), pp. 13613–13629.
- [15] Z. WANG, Y. WANG, AND Y. JIAO, *Uncertain multi-objective hazardous materials transport route planning considering resilience and low-carbon*, *IEEE Access*, 11 (2023), pp. 26921–26931.
- [16] L. XU, F. JIA, X. LIN, AND L. CHEN, *The role of technology in supply chain decarbonisation: Towards an integrated conceptual framework*, *Supply Chain Management: An International Journal*, 28 (2023), pp. 803–824.
- [17] V. YOLCU, M. YUCEL, AND D. AYDIN, *Solution of fiber raman amplifier model using binary search equation-based adaptive artificial bee colony algorithm*, *Optical Engineering*, 62 (2023), pp. 26105–26105.

*Edited by:* Zhenling Liu

*Special Issue on:* Cloud Computing for Intelligent Traffic Management and Control

*Received:* Mar 20, 2023

*Accepted:* Sep 1, 2023





## TIME WINDOW ORIENTED IOT VEHICLE PATHWAY STUDY FOR THE DYNAMICALLY CHANGING NEEDS OF E-COMMERCE CUSTOMERS

**Abstract.** The main dynamic truck routing problem also presents a significant difficulty in the logistics sector, which is an unavoidable development trend of the contemporary technological changing society. A dynamic vehicle routing problem with time window model is suggested by the study in order to establish an effective and low-energy dynamic response method. The fundamental concept is to disrupt the conventional strategy of static dynamic consumers responding in time slots by dividing the dynamic time window into a static time window with several time slice intervals. The study makes use of cutting-edge ideas including dynamic attitude, before-and-after time slicing, and continuous optimisation while proposing a new method for model solution to optimise dynamic vehicle route issues effectively and affordably. The study employs the Solomon optimisation dataset and runs simulation studies on the Java platform to confirm its efficacy. The experimental findings demonstrated that the optimisation technique employed in the study reduced the cost of travelling by 83.8 miles while also considerably increasing the average vehicle utilisation by 3.6%. Because driving distance cost and vehicle number cost are typically positively connected with dynamic attitude, the study employs solutions that can increase dynamic response efficiency and save money. As a result, their robustness is higher.

**Key words:** Internet of Things; path optimization; time windows; dynamic demand; cost

**1. Introduction.** All businesses are being moved towards intelligence by technological advancements, particularly by the emergence and growth of the Internet of Things (IoT) and intelligent algorithms, which are creating new prospects for the logistics sector. With the continuous development of science and technology, artificial intelligence has been elected as one of the hottest industries in recent years and is widely used in various industries. The field of intelligent vehicles has also introduced this advanced technology. Such transformation and upgrading plays an important role in national economic development and can greatly reduce the costs of vehicle driving. Many directions of its field have become research trends. In addition to automated driving, smart automobiles also effectively combine logistics and transport, which frequently leads to phenomena like bad customer experience because of fatal flaws like slow response in the traditional logistics industry, from the technical layer, industrial layer and application layer three aspects of the optimization industry, in the technical layer, will gradually change from manual to automation, to solve the problem of labor cost; At the industrial level, the transportation industry has become more network intelligent and performs better in route planning and other aspects. At the application level, transport vehicles are transformed into mobile intelligent terminals to provide convenience for people's life and work [2]. It is especially crucial to enhance the optimisation of dynamic vehicle path difficulties since task orders from IoT clients frequently alter in real time and with uncertainty in some medium and heavy transport operations. Vehicle scheduling is the process of determining the optimal route to take in order to satisfy various client needs and keep driving expenses, for example, within tolerable bounds [11]. Vehicle path optimisation (PO) systems with temporal windows (TW) are much more complicated. When it comes to the customer's dynamic demand (DD), the vehicle must first and foremost satisfy the customer's wants while attempting to design the optimum route to minimise driving expenses, etc., while still finishing the assignment on schedule [1]. Innovations in the logistics sector depend on research into putting forth a fresh set of optimisation models.

The first section of this research introduces the vehicle scheduling research application; the second section describes the construction process of this initial model and introduces the forbidden search algorithm as the model solving algorithm; the third section uses the Solomon dataset and conducts parameter tuning experiments on the forbidden algorithm to visualise the customer distribution; and finally, the novel model is compared with the traditional model. Finally, Section 4 presents a thorough analysis of the experimental data from the study and draws the conclusion that the Dynamic Vehicle Routing Problem with Time Window (DVRPTW) model

paired with the study's forbidding algorithm can be more effectively used to vehicle PO.

**2. Related works.** Adapting vehicle pathways to the DD of IoT clients is the key to attaining logistics vehicle scheduling in the logistics sector. To further improve the algorithm's learning capacity and combine user activity paths to produce optimal vehicle paths, Dai X et al. proposed a process model based on vehicle dynamic routes and user scenarios in order to realize the convenience of shared public transportation and reduce the burden on urban roads. According to the bus data set of multiple routes, the travel demand of the last kilometer is predicted, and the dynamic path is planned, and the shared public transportation is dynamically routing and scenario-based operation. Five prediction functions including time and location are used to further improve the learning ability of the algorithm, and the generation of the optimal vehicle path is realized by combining the user activity path [3]. According to Merka Z and colleagues, blockchain technology has a significant impact on the logistics sector, and the addition of IoT can make the entire logistics process more transparent, traceable, and productive. Supply chain refers to the whole process of producing parts, making intermediate products and final products, and finally sending them to consumers, involving suppliers, manufacturers, distributors and consumers through the connection of upstream and downstream members of the network chain structure. Supply chain management can minimize the entire system cost. They combined a number of cases to show the effectiveness of the current use of blockchain technology in a number of logistics industries, and they then proposed a method for building models to address the development limitations of the technology [8]. Lin R et al. proposed a logistics robot based on laser navigation system. The model consists of two sub-machines, which respectively contain two units driving the front and back, and both move in the same direction. The front-mounted laser is used for obstacle detection, and the on-board navigation system is used for positioning, while the rear laser is used for pallet picking [7]. Lasers at the front and rear ends are utilised for obstacle detection and logistics path planning, respectively. Wang X et al. proposed a vehicle PO model based on ant colony algorithm, introduced adaptive mechanism to tune the parameters, solved the shortcomings of the classical ant colony algorithm such as too fast convergence and too slow convergence, and achieved the effect of lowest cost as well as optimal path, and finally the study used data from two logistics enterprises in Huainan to verify the effectiveness of the heuristic intelligence algorithm.

The vehicle PO problem, which is one of the main challenges the system is currently facing, takes the maximisation of platform profit as the goal, while taking into account the driver's shift rest and other issues. Guo J et al. argued that the cross-city carpooling system needs to be further optimised, and the study proposes a VRP model that uses the sequential path construction method as the initial value solution path while introducing the variable neighbourhood search. The model was solved, and experiments were used to confirm its efficacy [4]. Himmich et al. investigated a novel primitive column construction framework applied to the vehicle scheduling problem and embedded it into a column generation scheme to solve the subproblem, and experimentally demonstrated that the method's solution falsity was reduced by a factor of seven in comparison to the conventional method by abandoning primitive algorithms like adjacency [5]. Vehicle scheduling was regarded by Li Y et al [6]. as an essential stage in managing public transit. In order to address the issue of randomness in vehicle assignment, the study applied the discrete artificial bee colony algorithm to the vehicle scheduling problem. It also used small differential coding and decoding to address the discretization defect and introduced three neighbourhood search schemes, such as initialization rules and heuristics, including discrete scheme, heuristic scheme, and learnable scheme to achieve the best vehicle scheduling optimisation results [5]. In order to find the best solution to the vehicle path problem, Raeesi R et al. proposed an optimisation of an electric vehicle path problem with TW and mobile battery swapping [12]. The study combined dynamic programming and integer programming algorithms.

Dynamic reaction models clearly play a crucial role in the logistics of vehicle transportation. This work proposes a dynamic vehicle PO model based on a taboo search algorithm that interprets the problem as a set of time-sliced intervals for a static vehicle problem and enhances performance, including cost.

**3. One-stage DD-based Vehicle Path Study with TW.** It is inevitable when customer demand changes dynamically in the process of logistics transportation. However, there are few researches on dynamic demand response at present, and most logistics still maintain the traditional fixed path planning scheme, which will greatly reduce the efficiency of vehicle transportation and increase its cost. Moreover, with the increasing number of logistics employees, it is more necessary for the model to realize the reasonable division of staff and

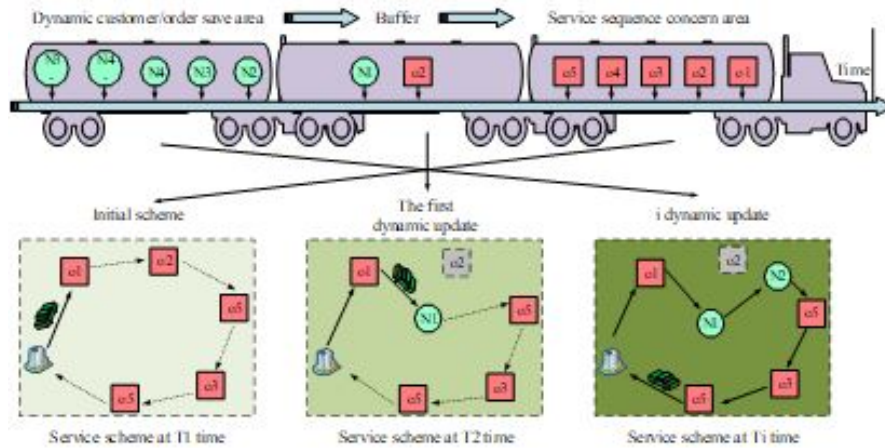


Fig. 3.1: Schematic diagram of DVRPTW

work area. However, due to the randomness of dynamic requirements and the diversity of application scenarios, it is usually difficult to construct dynamic models, so a systematic construction system is needed. The PO of a vehicle is the key to achieving efficient work in logistics, which in practical application scenarios is usually reflected in a dynamic response process, requiring a dynamic vehicle path problem model with TW.

**3.1. Vehicle Path Modeling for IoT Customer DD.** In the static vehicle scheduling model, the time window includes hard time window and soft time window. The soft time window uses the penalty function for vehicle path planning. This method not only cannot obtain the best path, but also relies on the selection of penalty factors to a great extent, which is highly subjective. The hard time window provides the customer access of the specified time window, which can get the shortest path and the observability is stronger. Therefore, the model chooses a hard time window. In the dynamic vehicle scheduling model, the constraint on the time window is usually abandoned for the sake of timeliness, but this will also cause lateness and reduce customer satisfaction. Therefore, the concept of continuous time slice and equivalent time window transformation is introduced to optimize the time window. At the same time, dynamic attitude is introduced to simplify and improve the model. There is a certain relationship between dynamic attitude and time slice. The direction of commodity flow reflects the chain in which the logistics transport is located, including the three steps of the front and middle end. When it is at the front and middle end of the process, the vehicle needs to meet the requirements of the IoT customer, collect the goods in order and subsequently deliver them uniformly to the collection point. Online shopping is a common IoT model, individual users (IoT users) use smart devices to make online orders, followed by cooperation between merchants and logistics companies to achieve product pickup and transportation, in the process, the product transportation status will change in real time according to the DD of IoT users, which is the dynamic vehicle scheduling (Dynamic vehicle scheduling Problem (DVSP) process [14]. The DVRPTW is shown in Figure 3.1.

IoT users are volatile in terms of when and where they place their orders, and therefore need to meet higher standards in terms of specific services TW. The study establishes the DVRPTW model with the aim of accomplishing more efficient IoT platform services at a smaller cost. As can be seen from figure 3.1, vehicles are subject to the constraints of the undirected connectivity network graph, denoted by  $G(Node, Edge)$ , which means that the correct edge *Edge* of the full user *Node* is selected. the dynamic vehicle path problem with TW is characterised by the inclusion of the length of the working day in the calculation process, where those who place an order before 1/2 working day belong to dynamic users, and vice versa, they are classified as static users on the next working day, which working day with priority processing rights [13]. The upfront improvement of the DVRPTW is similar to the static problem, with the expression of the primary and secondary optimisation

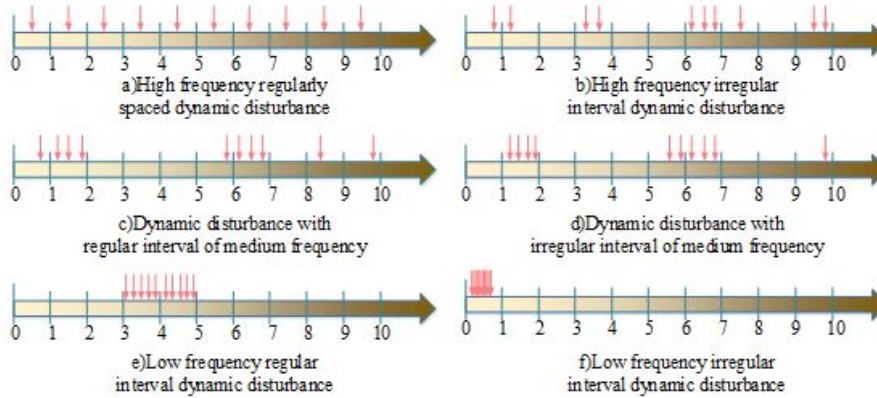


Fig. 3.2: Different dynamic event disturbance scenarios

objective function shown in Equation 3.1.

$$\begin{cases} \text{Primary.objective} = \min \sum \sum \sum x_{[i],[j],[k]} \\ \text{Secondary.objective} = \min \sum \sum \sum d_{[i],[j]} x_{[i],[j],[k]} \end{cases} \quad (3.1)$$

In Equation 3.1,  $d_{[i],[j]}$  is the Euclidean distance between the random user  $i$  and  $j$ ;  $x_{[i],[j],[k]}$  represents the  $(0, 1)$  decision variable of whether the vehicle passes through two customers. The two parameters are calculated as shown in Equation 3.2.

$$\begin{cases} d_{[i],[j]} = \sqrt{(i_x - j_x)^2 + (i_y - j_y)^2} \\ x_{[i],[j],[k]} = \begin{cases} 1, \text{ if } i - j \text{ are connected} \\ 0, \text{ if } i - j \text{ aren't connected} \end{cases} \end{cases} \quad (3.2)$$

When an IoT user places an order before half a working day, it is already part of the pre-logistics phase, when a combination of dynamic and static customer status is required for real-time vehicle logistics updates. Assuming the vehicle state is on the way to the next user, the next logistics task will not be available until after that task state has been completed. And it needs to return to the initial point after all tasks have been completed, while the loading volume needs to meet the requirement of less than the maximum amount [16]. The difficulty of running the model is related to the value of its dynamic attitude, which refers to the ratio of dynamic users to the total number of users in a fixed time horizon, as shown in Equation 3.3.

$$Dyn = \frac{N_c^d}{N_c^d + N_c^s} \quad (3.3)$$

In the Equation 3.3,  $N_c^d$  is the total number of dynamic users;  $N_c^s$  is the total number of static quantities. The equation is simple and easy to calculate, but when the amount of time and the dynamic/static number of users are all 10, there are six different forms of its dynamic attitude, and the performance can produce limitations, as shown in Figure 3.2.

As can be seen from Figure 3.2, the neighbourhood search first generates an initial solution, followed by the selection of a neighbourhood range, or local search range, according to the definition. The neighbourhood solutions are compared one by one by evaluation, and the set of filtered neighbourhood solutions is shifted, and finally the local optimum is recorded, and by cycling through the above steps, the global optimum is eventually

ranked. The amnesty rule can be used to solve the problem of falling into a local optimum, when a candidate value has violated the taboo rule, it needs to be released by this principle, and the release principle enables the restricted element to obtain a more optimal value. Among them, the indicators of the taboo table are mainly the taboo object and the taboo length, and the taboo length refers to the number of iterative failures. ts may obtain both correct and infeasible solutions throughout the search process, the infeasible solution has a positive effect on the global optimal search, so it can bring great convenience to the algorithm operation [10]. Therefore, the study introduces an adaptive penalty function and proposes an adaptive adjustment parameter as an optimisation. The adaptive penalty function is calculated as shown in Equation 3.4.

$$Dyn = \frac{N_c^d + DN(T/2n_{ts})}{N_c^d + N_c^s + N(T/2n_{ts})} \quad (3.4)$$

In Equation 3.4,  $N(T/2n_{ts})$  is the amount of time slices in the full process;  $DN(T/2n_{ts})$  is the amount of time slices in the request made by the dynamic user. This equation incorporates all the factors that will have an impact on the performance of the dynamic attitude. To improve the dynamic scheduling process of the vehicle path, multiple time slices of are set for the total duration of half a working day, and the initial operating moment is set to . The expression for half a working cycle is shown in Equation 3.5.

$$T_{lim} = [(t_0, t_0 + n_{ts}), (t_0 + n_{ts}, t_0 + 2n_{ts}), \dots, (t_0 + nn_{ts}, T/2)] \quad (3.5)$$

In Equation 3.5,  $T_{lim}$  denotes half a working day. It can be seen that the time slices are not discrete. According to the setting of the continuous interval, the study defaults that dynamic customers in each time slot are interpolated to the corresponding vehicle paths, and eventually real time updates of the paths are achieved. The model responds to the dynamic customer problem when it determines that the vehicle can achieve the user's time requirements, and the determination rules are shown in Equation 3.6.

$$f(t) = \begin{cases} 0, & b_i < t_0 + in_{ts} \\ 1, & b_i > t_0 + in_{ts} \end{cases} \quad (3.6)$$

The dynamic user demand TW is known to be denoted as  $[a_i, b_i]$ . In the equation 3.6, when  $b_i$  satisfies the condition greater than  $t_0 + in_{ts}$ , then the system is able to demand the corresponding user demand. Conversely, the vehicle system will reject the corresponding user demand. When the vehicle successfully corresponding demand, the fastest time for the dynamic user to complete the service is  $t_0 + in_{ts}$ . The updated TW expression is shown in Equation 3.7.

$$a_{i\_r} = \begin{cases} t_0 + in_{ts}, & a_i < t_0 + in_{ts}, b_i > t_0 \\ a_i, & a_i > t_0 + in_{ts}, b_i > t_0 \end{cases} \quad (3.7)$$

To ensure that the vehicle path can be as smooth as possible to facilitate vehicle travel and to further save user time, the study introduces an equivalent TW transformation concept that uses constraints to update the user task completion status as shown in Equation 3.8 [15].

$$[a_i, b_i] = \begin{cases} a_i = INT(t_{ik\_a}) \\ b_i = INT(t_{ik\_a}) + 1 \end{cases} \quad (3.8)$$

In the Equation 3.8,  $t_{ik\_a}$  denotes the real moment when the vehicle numbered  $k$  completes the user  $i$ ;  $INT$  denotes the rectification function. At any time on the node, when the first time need to serve the customer, will use the new vehicle, then the task completion time is the moment of arrival of the new vehicle. In the workflow, maintaining the customer's time rights comes first, and equation 3.8 allows for the optimal selection of paths. When the vehicle path is subject to this restriction, then the vehicle path is said to be unique. When the path to task completion is broken, then path planning violates the TW constraint, and conversely, the vehicle is able to complete the task successfully and on time. The study transforms the dynamic vehicle path problem into a

static vehicle path problem for multiple short time intervals by means of dimensionality reduction at the time level. The basic idea is to refine the path at the next moment in the context of the previous time slice under the condition of minimum driving cost, and to keep cycling through the process until the end of the current working day, which is expressed as shown in Equation 3.9.

$$\begin{cases} \min F(y + [s + 1]|y_{[s]}), s = 1, 2, \dots, t_{ns} + 1 \\ y_{[s]} = \min \sum \sum \sum d_{[i_{-s}][j_{-s}]} x_{[i_{-s}][j_{-s}][k_{-s}]} \end{cases} \quad (3.9)$$

In equation 3.9,  $s$  denotes the time slice;  $F(y + [s + 1]|y_{[s]})$  denotes the driving cost. This equation is the objective function of the DVRPTW model. In the DVRPTW model, the driving path constraint of the vehicle is shown in equation 3.10.

$$\begin{cases} \sum_{k \in K} \sum_{j \in \Delta^*(i)} x_{[i_{-s}][j_{-s}][k_{-s}]} = 1, \forall i \in N \\ \sum_{j_{-s} \in \Delta^+(1)} x_{[1][j_{-s}][k_{-s}]} = 1, \forall k_{-s} \in K \\ \sum_{i_{-s} \in \Delta^-(j-s)} x_{[i_{-s}][j_{-s}][k_{-s}]} - \sum_{i_{-s} \in \Delta^+(j-s)} x_{[i_{-s}][j_{-s}][k_{-s}]} = 0, \forall k_{-s} \in K \forall j_{-s} \in J \\ \sum_{i_{-s} \in \Delta^-(n+1)} x_{[i_{-s}][n+1_{-s}][k_{-s}]} = 1, \forall k_{-s} \in K \end{cases} \quad (3.10)$$

In equation 3.10, this represents the constraint that the customer can only be served once; the vehicle departure point constraint; the vehicle departure place constraint; and the vehicle return constraint, i.e., the need to revert to the initial departure point with expression . The constraint on vehicle loading is calculated as shown in Equation 3.11.

$$\sum_{i_s \in N} d_{[i_s]} \sum_{j_s \in \Delta^*(i)} x_{[i_s][j_s][k_s]} \leq C, \quad \forall k_s \in K \quad (3.11)$$

As there is a rated load requirement for the vehicle, the DVRPTW model needs to constrain the maximum loading for each time slice, i.e., its load cannot be greater than the rated load size.

**3.2. Improvement of Tabu Search Algorithm Based on Dynamic PO and Other Policies.** The study introduces the Tabu Search (TS) algorithm, an algorithm that incorporates the concepts of taboo tables and domain search. It sets the storage rules as agile taboo criteria, and when the computation is constrained by the taboo length as well as the amnesty rules, the search process can be carried out more efficiently and accurately, resulting in the best value for the global search. TS is widely used in similar fields such as VRPTW, where neighbourhood search is an extremely important part of the algorithm. The underlying idea is that an iterative process searches the neighbourhood of the current value and eventually finds the optimal solution. When the neighbourhood range is too large, it will increase the burden of running the algorithm and cost more time, but at the same time, the local best value will be obtained [9]. The most common current strategies are single-path based internal search (*Or - opt*, *2 - opt*) and dual-path (*2 - opt\**, *swap/shift*) based search, respectively, for a total of four neighbourhood search strategies, expressed as shown in Equation 3.12 [18].

$$S_1 \rightarrow \text{Selection function} \begin{bmatrix} Or - opt \\ 2 - opt \\ 2 - opt^* \\ Swap/shift \end{bmatrix} \rightarrow S_2 \quad (3.12)$$

In the equation 3.12,  $S_1$  denotes the initial solution of the vehicle route. denotes the final solution of the neighbourhood search. Where, AAA is constructed as shown in equation 3.13.

$$S_1 = \{0, c_1, c_2, c_3, c_4, \dots, c_n, n + 1\} \quad (3.13)$$

In the equation 3.13,  $c_n$  denotes the number of iterations for  $n$  time slices. Firstly the vehicle must meet the rated capacity as well as the TW condition, then nodes of completed users and newly added customers are arbitrarily selected and the former is placed to any node to achieve dynamic update. The neighbourhood search process is shown in Figure 3.3.

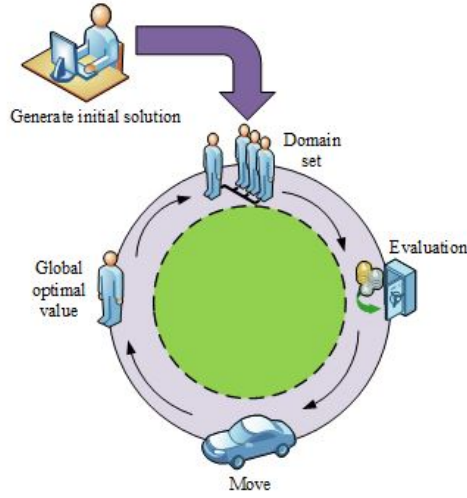


Fig. 3.3: Flow chart of neighborhood search strategy

As can be seen from Figure 3.3, the neighbourhood search first generates an initial solution, followed by the selection of a neighbourhood range, or local search range, according to the definition. The neighbourhood solutions are compared one by one by evaluation, and the set of filtered neighbourhood solutions is shifted, and finally the local optimum is recorded, and by cycling through the above steps, the global optimum is eventually ranked. The amnesty rule can be used to solve the problem of falling into a local optimum, when a candidate value has violated the taboo rule, it needs to be released by this principle, and the release principle enables the restricted element to obtain a more optimal value. Among them, the indicators of the taboo table are mainly the taboo object and the taboo length, and the taboo length refers to the number of iterative failures. It may obtain both correct and infeasible solutions throughout the search process, the infeasible solution has a positive effect on the global optimal search, so it can bring great convenience to the algorithm operation [10]. Therefore, the study introduces an adaptive penalty function and proposes an adaptive adjustment parameter as an optimisation. The adaptive penalty function is calculated as shown in Equation 3.14.

$$f(x) = y_{[s]} + \alpha Q_{over} + \beta T_{over} \quad (3.14)$$

In equation 3.14, the adaptive penalty function includes three cost indicators for distance, overload and violation time, with the latter two cost parameters,  $\alpha$  and  $\beta$ , relating to the adaptive regulation parameters as shown in equation 3.15 [17].

$$\begin{cases} \alpha \begin{cases} \alpha/(1+\theta), & \text{if } Q_{over} = 0 \& \alpha \geq \alpha_1 \\ \alpha \cdot (1+\theta), & \text{if } Q_{over} = 0 \& \alpha \leq \alpha_1 \end{cases} \\ \beta \begin{cases} \beta/(1+\theta), & \text{if } T_{over} = 0 \& \beta \geq \beta_1 \\ \beta \cdot (1+\theta), & \text{if } T_{over} = 0 \& \beta \leq \beta_1 \end{cases} \end{cases} \quad (3.15)$$

In equation 3.15,  $\alpha$  and  $\beta$  correspond to the vehicle load as well as the time cost indicators of the adaptive penalty function, respectively. It can be seen that  $\alpha$  varies with the total user demand, and the process is controlled by the adaptive regulation parameter. When the total demand of the user is greater than the maximum load of a vehicle, the adaptive parameters will be appropriately increased, so that the value of the fitness function can be increased; In the opposite case, an appropriately reduced value is needed to achieve

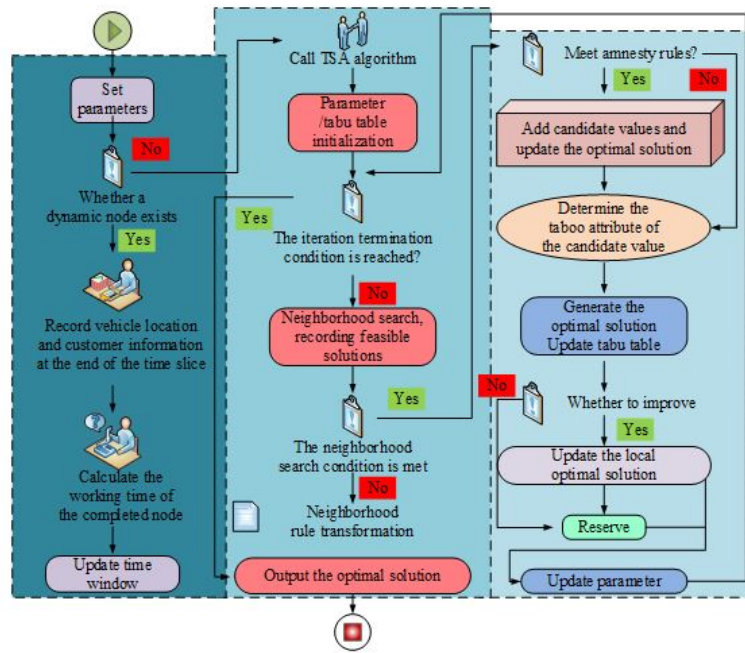


Fig. 3.4: Flowchart of Tabu algorithm implementation

a reduction in the fitness function. Moreover, if a vehicle completes its task overtime,  $\beta$  will also be adjusted accordingly by parameter  $\theta$ , realising the automatic adjustment of the adaptation function. That is, when the vehicle does not arrive at the customer point on time, the adaptive parameter will increase it to achieve the improvement of the fitness function, otherwise the condition will automatically be lower [20]. In summary, the flow chart of the TS algorithm is shown in Figure 3.4.

The algorithm needs to ensure the minimum damage degree of the traversed path, which is realized by the adaptive penalty function. The concept of equivalent time conversion is used to achieve the purpose of partial path validity. For example, if a customer's time slice has been completed, a time window constraint is required to complete the correction. The generation of the initial solution is also assisted by the time window constraint. Customers meeting the constraint are randomly selected and inserted into the corresponding vehicle chain. The process is repeated continuously. When the number of customers reaches the capacity constraint, the next vehicle chain can be added. Among them, the adaptive penalty function optimizes the fitness value by adjusting its adaptive parameters, and then realizes the cost control. Initially, the vehicle traverses all customers at random time slices with the path planning direction, entering completed users at the end of the time interval; then, the static and dynamic users recorded by the system generate a new undirected connectivity network graph; finally, the system introduces TS to solve the VRPTW problem, and based on the optimal solution result, the path is updated in real time and provided to each vehicle.

**4. Experiments on the Performance of a TS-based System for Optimizing Dynamic Vehicle Path Problems for IoT Users with TW.** As the dynamic vehicle scheduling model adopted in this study has time window constraints, the performance simulation experiments of this model mostly adopt Solomon optimization data set, because the data set introduces time slice rules, and each time slice is equivalent to static vehicle path optimization, which is more suitable for the study of practical models. Therefore the Solomon optimisation dataset was selected as the basis for the study, and the optimisation scheme was specified as a random allocation of the next 50 users in the form of  $\max(0, e_i - \theta dis_{oi} - r)$  if the first 50 users all appeared at moment 0. The experimental environment parameters, i.e., the basic characteristic information of each data



in the dataset, are shown in Table 4.1.

Table 4.1: Experimental environment parameters and data set characteristics

Content	Parameter
Computer model	MacBook Air 13.3 Core i5
CPU	1.8GHz
Memory	8G
SSD	512G
System	Windows 10 64bit
Programming environment	Java JDK-8u251
	R1 R2 C1 C2 RC1 RC2
Number	12 11 9 8 8 8
Working day length	230 1000 1235 3390 240 960
Customer service hours	10 10 90 90 10 10
Total demand	1457 1457 1810 1810 1725 1725
Vehicle capacity	200 1000 200 700 200 1000
Theoretical minimum vehicle	7.28 1.47 8.63 2.58 8.63 1.73
Mean time window	86.97 453.75 85.42 920.52 85.44 369.77
Mean EDD	0.13 0.13 0.17 0.18 0.17 0.13
Mean LoU	0.52 0.46 0.53 0.58 0.53 0.50

It is known that the study divides the DVRPTW model in the form of time slices and analogises each time slice to a refinement process for a static vehicle path problem. The study began by subjecting each time slice to performance validation experiments. In Table 4.1, information on the 56 data features with temporal constraints in the Solomon optimisation dataset is included.

**4.1. Parameter Optimisation Experiments and Performance Testing of the TS.** To investigate the relationship between parameter selection and algorithm optimisation results, a significance hypothesis testing method was introduced. The first group was set up as control group A/B with different parameters  $\alpha/\beta$  and the second group was control group C/D with different adaptive parameters. After conducting several experiments on the data from the two groups respectively, a two-sample test was finally introduced for variance analysis and the results are shown in Figure 4.1.

The experiment was set up with values for the  $\alpha/\beta$  parameters of (0.1/0.2) and (0.9/1.0) for groups A/B, respectively. From Figure 3.1(a), it can be seen that the experimental values obtained in Group A are significantly higher than those in Group B. The mean value of Group A is 1786.93, while that of Group B is only 1756.42, a difference of 30.51. The median values of the two groups are 1785.98 and 1736.27 respectively. The standard error and standard error of the mean for group A were 58.41 and 18.47 respectively, while those for group B were 74.91 and 23.70 respectively. It can be seen that the SEM value for group A was 5.23 smaller than that of group B, indicating that group B has better dispersion and stability of the mean. According to the significance hypothesis test, when the confidence level is set to 0.05, there is no significant difference between the resultant values under both conditions of assumed unity of variance and assumed unity of variance, so the algorithm is robust to the two parameters. In summary, the study selected data set B as the value for the  $\alpha/\beta$  parameter. Figure 4.1(b) shows the control experiment set up to compare the values of the adaptive parameters, where the values for groups C/D are 0.5 and 0.8 respectively. It can be seen that the test means for groups C/D are 1785.91 and 1728.19 respectively; the SD and SEM values for the two data sets are (58.42/18.48) and (50.98/16.12) respectively. Again, the two sets of data were tested for significance at a confidence level of 0.05. The results showed that the test exhibited significant differences in both cases where the assumptions were flush and where the assumptions were not flush. Therefore, the value has a greater impact on the performance of the algorithm. The study carried out further tuning and optimisation of this parameter and set the value

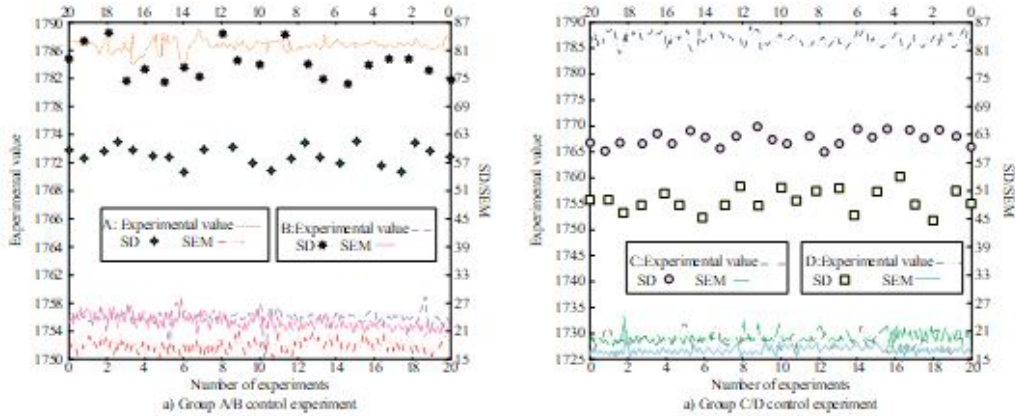


Fig. 4.1: Parameter tuning experiment

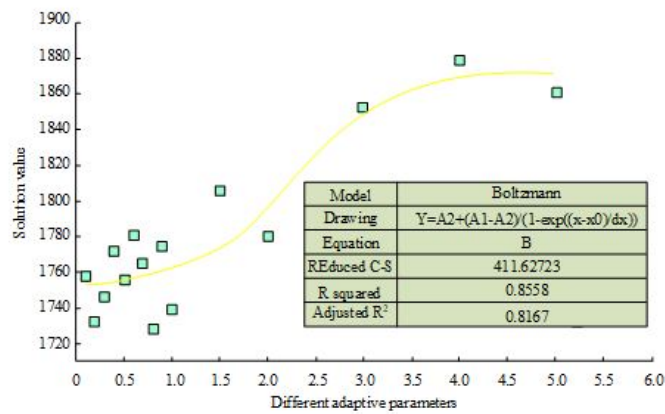


Fig. 4.2: Fitting results of adaptive parameters

conditions to  $[0 \ 0.11]$  and  $[1 \ 1 \ 5]$ . After 10 iterations, the experiments took the mean value of the data as the final selected value and used the Boltzmann function for scatter fit analysis. Figure 4.2 shows the fitting results and the final parameter selection values. As can be seen in Figure 4.2, the optimisation results of the algorithm show a tendency to become progressively weaker as the value of the adaptive parameter rises, with the algorithm optimising best when it takes values within 1, and the algorithm performance deteriorating when it is in the (1, 6) interval. Therefore, the study uses the average minimum value in the (0, 1) interval, i.e.  $\theta = 0.8$ . In general, the above experiments prove that the selection of the three parameters will change the total demand. Among them, the adaptive parameters are positively correlated with the optimization results, but the optimization effect will deteriorate when a certain equilibrium point is reached.

In Figure 4.3(b), the Z-axis represents the customer demand and the X/Y-axis represents the location of the customer. When the time slice interval is certain, if the dynamic customer's minimum order demand time lies at the end of the area, the corresponding service time will not be less than the end of the interval. If the customer's demand time does not meet the requirement, the system will automatically reject the request. In the

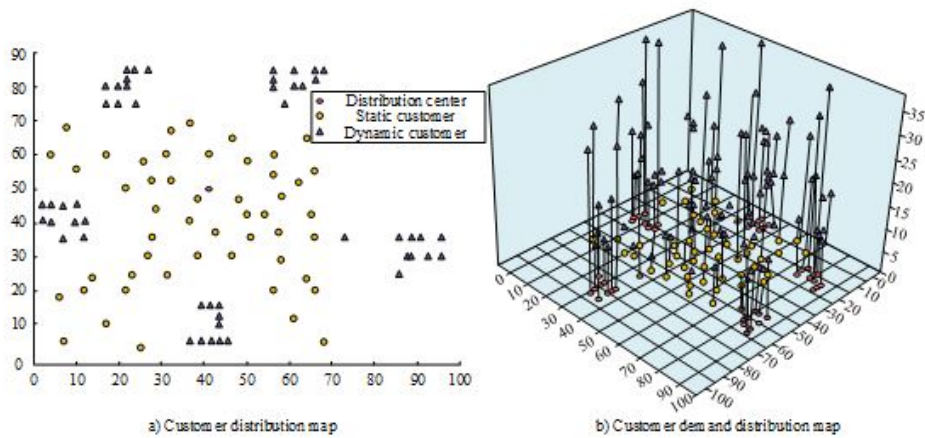


Fig. 4.3: Customer classification and customer demand analysis

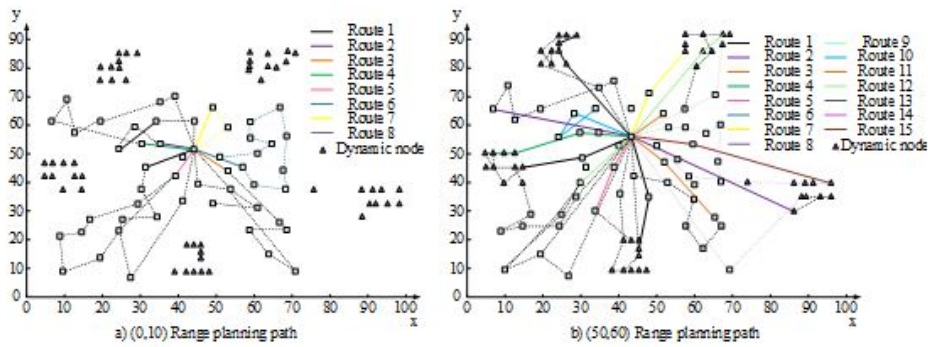


Fig. 4.4: Dynamic vehicle routing optimization simulation at different time slices

current prevalence of online shopping, customer points, points of sale and service points are basically interactive through intelligent means. Among them, the dynamic change of the relationship between the point of sale and the point of service requires the DVRPTW model to realize. After the vehicle gets the customer’s demand, it should respond to it in a timely manner. The model can visualize static users and dynamic users, which can improve the operation efficiency of the model. First of all, static users should be satisfied first, because the dynamic requirements of users will affect the path generation, while the optimal path of static users is fixed, which greatly reduces the operation cost. Dynamic user paths follow static user paths until static user requirements are fully addressed. Therefore, the visualization of user distribution is very necessary. The study selects Instance RC102 as the test set and also sets the time slice parameter  $n_{ts}$  to 10, which yields 50 dynamic users in the first time slice interval. The study did dynamic vehicle PO simulations for the time slice intervals of  $[0, 10]$  and  $[50, 60]$  respectively, and the results are shown in Figure 4.4.

In figure 4.4, the dotted line represents path planning that does not incorporate the needs of dynamic customers; realisations represent path planning that incorporates the needs of dynamic customers. The purple symbols then indicate dynamic customers. At the initial moment, the merchant assigns eight vehicles to serve

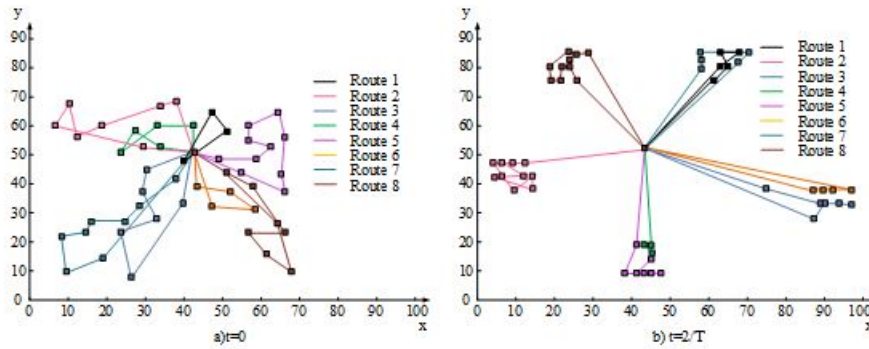


Fig. 4.5: Path analysis without considering dynamic customers

the needs of 50 static customers from the previous working day, and the distance cost of driving is calculated as 743.81. Once the system has completed 50 dynamic customers, its distance cost rises to 1504.53 and the number of vehicles rises to 15. PO performance indicators include both vehicle load factor as well as time utilisation, with time and vehicle utilisation reaching 100% and 96.51% respectively, making the DVRPTW model extremely efficient for optimisation. However, the customer's TW constraint causes some waste of time, as the vehicle has already departed from the work centre at this point and its needs to complete its first order before the dynamic user's task can be considered. There are a large number of constraint algorithms in the model, which results in the vehicle having to complete the specified task before it can start the next order task. To further understand the impact of different conditions on the performance of the DVRPTW model, the study compares it to a traditional PO. The traditional approach does not consider requests from dynamic customers in the first half of the working day, serving only static users from the previous working day, and only processes dynamic customer requests from the previous half of the working day in the second half of the working day. The traditional optimisation path is shown in Figure 4.5.

As can be seen from Figure 4.5, the conventional PO divides it into two time slots, with the total number of vehicles in both time slots being 8. In the first time slice interval, the distance cost of the vehicles is 744.38. The distance cost of travelling in the second time slice interval is 843.95, and the total distance cost is 1588.33, which is more than 83.8 compared to the DVRPTW model. based on further calculations, the average vehicle utilisation for the two time slots reached 47.13% and 60.62% respectively, with a total average utilisation of 53.875%, which is significantly lower than the average load factor of 57.48% for the DVRPTW model. The experimental results show that the new DVRPTW optimisation model, which allows for better cost-efficiency, works better than traditional optimisation methods in terms of response speed and vehicle utilisation. Due to the excessive amount of data,  $T/2$  time slice in the research model was selected for comparison. The performance of each model in terms of the number of customers, actual vehicle load, vehicle load utilization rate and route cost is shown in Table 4.2.

As can be seen from Table 4.2 above, compared with the traditional model, the average vehicle utilization rate of the optimized model has increased by 35.87%. This is because after route optimization, the number of customers that vehicles can satisfy at the same time has increased, which has greatly improved the vehicle loading rate and made it realize the maximum load on one route. The total customer task completion is twice that of the traditional model. Although the path cost of the optimization model is higher, it achieves a substantial improvement in completion efficiency at the lowest possible cost. This is because the path planning consistency of the optimized model is significantly better, and the vehicle utilization rate is naturally higher, which also leads to the rest of the performance improvement. The study repeated the above experiments several times and recorded the distance cost and vehicle cost for the two time slice intervals respectively, distributed in a descending form, to obtain the relationship between driving cost and vehicle utilisation cost as shown in

Table 4.2: Performance comparison of each model

Model	Path number	Customer quantity	Actual load	Path cost	Loading rate/%
Optimization model	1	9	171	107.95	85.50
	2	6	129	110.05	64.50
	3	8	124	98.81	62.00
	4	9	149	115.02	74.50
	5	5	99	82.80	49.50
	6	6	69	76.15	34.50
	7	7	87	99.02	43.50
	8	8	147	113.07	73.50
	9	5	86	129.23	43.00
	10	4	61	60.96	30.50
	11	8	152	117.48	76.00
	12	6	146	93.23	73.00
	13	11	193	113.02	96.50
	14	6	95	155.14	47.50
	15	2	16	32.57	8.00
Sum total	100	1724	1504.52	96.50	
Traditional model	3	6	100	123.45	50.00
	4	3	60	78.73	30.00
	5	7	140	106.94	70.00
	6	4	80	114.45	40.00
	7	5	70	95.74	35.00
8	10	190	105.96	95.00	
Sum total	50	970	843.98	60.63	

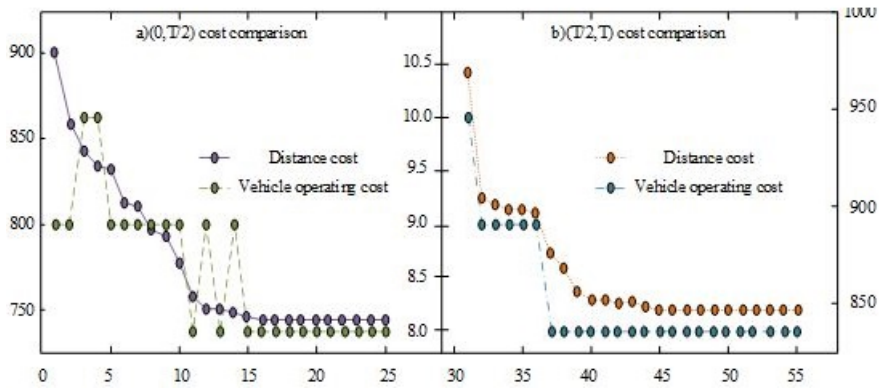


Fig. 4.6: Relationship between driving cost and vehicle use cost

Figure 4.6. As can be seen from Figure 4.6, the trends of vehicle use cost and driving cost are more or less the same in the two time slice intervals, showing a positive correlation, especially in the  $(2/T, T)$  time slice interval, where both costs decrease almost simultaneously, while in the  $(0, 2/T)$  time slice interval, the vehicle use cost changes more curvilinearly and the driving distance cost is relatively moderate. As can be seen from the  $(10, 15)$  interval in Figure 4.6(a), an appropriate increase in the number of vehicles may lead to a reduction in driving costs when the vehicle path is in the pre-fluctuation phase. Experiments were introduced to the

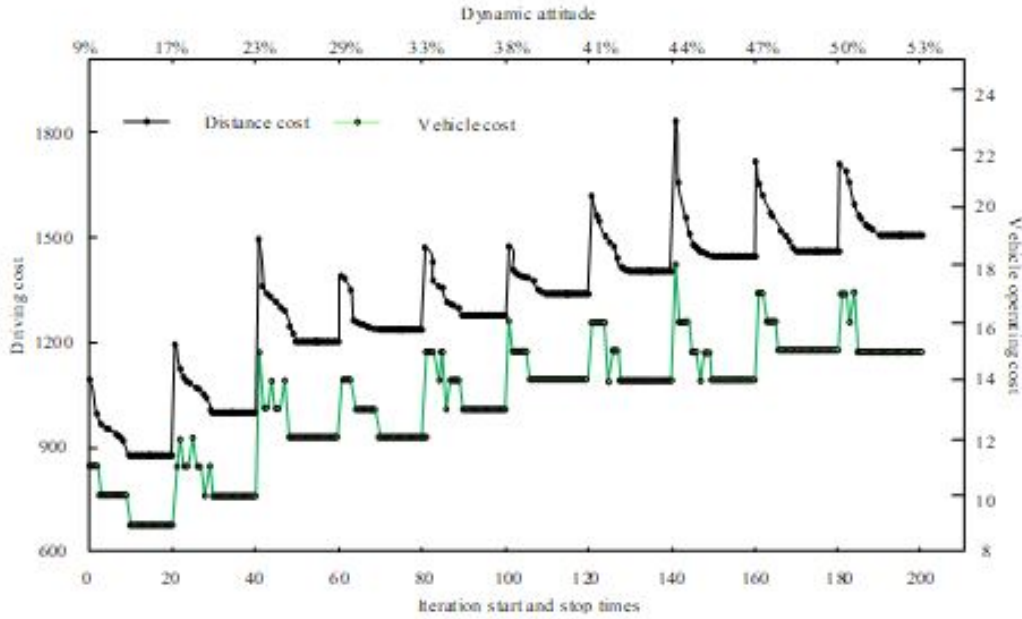


Fig. 4.7: Sensitivity of dynamic attitude to cost

InstanceRC102-50-50 case to explore the sensitivity of dynamic attitude to cost, and the results are shown in Figure 4.7. When  $t = 0$ , the total number of static customers is 50, and as the time slice increases, the number of dynamic users increases by 5 in turn. It can be seen that cost and dynamic attitude are roughly positively correlated. In particular, the correlation of the cost of vehicle use shows interval intermittency, and within a certain range, its cost correlation will show a decreasing or stable constant trend. The experimental results show that the system is more robust to DD when the rated capacity of the vehicle is larger. Until an optimal solution is chosen, its optimisation scheme revolves around the minimisation of the number of vehicles, but may produce a phenomenon where the two costs appear in an opposing relationship. According to the relationship between vehicle routing cost and use cost, it can be seen that the optimal balance value can minimize the cost of both sides, which indicates that the relationship between the two is proportional. The dynamic attitude is the comprehensive consideration of the number of dynamic customers, the time point of occurrence and the frequency of occurrence, which are all factors that lead to cost changes. In the early stage, the change of the number of customers will only affect the distance cost, but with the increase of the vehicle capacity, the ability to meet the dynamic demand is greater. Therefore, the cost is positively correlated with the dynamic attitude.

**5. Conclusion.** The dynamic vehicle path problem must be optimised since inadequate merchant response frequently results in a bad IoT consumer experience. In the paper, a DVRPTW model is suggested, one is modelled, one is examined, and one is solved using a prohibited search method. Finally, the study utilised the Solomon optimisation dataset as the basis for the simulation experiments. Since the  $\alpha/\beta/\theta$  parameters in TS have a large impact on the performance of the algorithm, the study first conducted parameter tuning experiments on them using significance hypothesis validation. Through two different sets of control experiments, the  $\alpha/\beta/\theta$  parameters of 0.9/1.0/0.8 respectively were finally selected. The new optimisation approach and the conventional optimisation strategy were further compared in the experiments. The experimental results showed that the DVRPTW model achieved an average vehicle utilisation rate of 57.48% compared to only 53.875% in the traditional optimisation strategy, a relative decrease of 3.6%. This was in contrast to the traditional

optimisation strategy, which only achieved an average vehicle utilisation rate of 53.875%. Additionally, the trials were run to visualise consumer distribution and examine how dynamic attitude and cost interacted, yielding a generally favourable association between the two. The DVRPTW optimisation model, which is a better option for dynamic vehicle routing issues since it allows for faster dynamic reaction and cost reduction than conventional PO approaches, is proven to be reliable by all of the aforementioned trials. Though there are still some problems to be solved, for example, is not high enough and needs to be further optimized. In addition, this study only designed the dynamic response of the vehicle to the customer, without considering the customer's response to the vehicle position. Only when there is two-way interaction between the two, can the model achieve the highest operating efficiency.

## REFERENCES

- [1] M. ABDIRAD, K. KRISHNAN, D. GUPTA, AND X. LI., *A two-stage metaheuristic algorithm for the dynamic vehicle routing problem in industry 4.0 approach*, *Journal of Management Analysis*, 8 (2021), pp. 69–83.
- [2] D. AMBROSINO AND M. PAOLUCCI, *Special issue of networks: New network models and approaches for logistics and transportation*, *Networks*, 78 (2021), pp. 227–228.
- [3] X. DAI, M. CHEN, AND Y. ZHOU, *Optimal logistics transportation and route planning based on fpga processor real-time system and machine learning*, *Microprocessors and Microsystems*, 80 (2021), p. 107.
- [4] J. GUO, J. LONG, X. XU, AND M. Y. YUAN, *The vehicle routing problem of intercity ride-sharing between two cities*, *Transportation Research Part B: Methodological*, 158 (2022), pp. 113–139.
- [5] I. HIMMICH, I. E. HALLAOUI, AND F. SOUMIS, *Primal column generation framework for vehicle and crew scheduling problems*, *Networks*, 75 (2020), pp. 238–252.
- [6] Y. LI, Y. SHEN, AND J. LI, *A discrete artificial bee colony algorithm for stochastic vehicle scheduling*, *Complex System Modeling and Simulation*, 2 (2022), pp. 238–252.
- [7] R. LIN, H. HUANG, AND M. LI, *An automated guided logistics robot for pallet transportation*, *Assembly Automation*, 41 (2020), pp. 45–54.
- [8] Z. MERKA, D. PERKOV, AND V. BONIN, *The significance of blockchain technology in digital transformation of logistics and transportation*, *International Journal of E-Services and Mobile Applications*, 12 (2020), pp. 1–20.
- [9] O. B. OLMEZ, C. GULTEKIN, B. BALCIK, AND A. E. OZENER, *A variable neighborhood search based math heuristic for a waste cooking oil collection network design problem*, *European Journal of Operational Research*, 302 (2022), pp. 187–202.
- [10] A. PONMALAR AND V. DHANAKOTI, *Hybrid whale tabu algorithm optimized convolutional neural network architecture for intrusion detection in big data*, *Concurrency and computation: practice and experience*, 34 (2022), pp. 7038–7052.
- [11] S. POTTEL AND A. GOEL, *Scheduling activities with time-dependent durations and resource consumptions*, *European Journal of Operational Research*, 301 (2022), pp. 445–457.
- [12] R. RAESI AND Z. KG., *The electric vehicle routing problem with time windows and synchronised mobile battery swapping*, *Transportation Research Part B Methodological*, 140 (2020), pp. 101–129.
- [13] ———, *Coordinated routing of electric commercial vehicles with intra-route recharging and en-route battery swapping*, *European Journal of Operational Research*, 301 (2022), pp. 82–109.
- [14] I. SBAI AND S. KRICHEN, *A real-time decision support system for big data analytic: A case of dynamic vehicle routing problems*, *Procedia Computer Science*, 176 (2020), pp. 938–947.
- [15] H. M. STELLINGWERF, G. LHC, G. LAPORTE, AND A. K. B. BEHDANI, *The quality-driven vehicle routing problem: Model and application to a case of cooperative logistics*, *International Journal of Production Economics*, 231 (2021), p. 109.
- [16] S. YANG, L. NING, P. SHANG, AND L. C. TONG, *Augmented lagrangian relaxation approach for logistics vehicle routing problem with mixed backhauls and time windows*, *Transportation Research Part E: Logistics and Transportation Review*, 135 (2020), p. 101.
- [17] X. ZHANG, L. CHEN, M. GENDREAU, AND A. LANGEVIN, *A branch-and-cut algorithm for the vehicle routing problem with two-dimensional loading constraints*, *European Journal of Operational Research*, 302 (2022), pp. 259–269.
- [18] Y. ZHAO, B. DU, AND S. LI, *Agile satellite mission planning via task clustering and double-layer tabu algorithm*, *Computer Modeling in Engineering and Science*, 122 (2020), pp. 235–257.

*Edited by:* Zhenling Liu

*Special Issue on:* Cloud Computing for Intelligent Traffic Management and Control

*Received:* Jun 26, 2023

*Accepted:* Aug 1, 2023







## PROTECTING DATA AND PRIVACY: CLOUD-BASED SOLUTIONS FOR INTELLIGENT TRANSPORTATION APPLICATIONS

**Abstract.** The interaction between transportation networks and intelligent transportation systems has been revolutionized by cloud computing. However, the reliance on cloud-based solutions raises security and privacy concerns. This article examines the challenges of safeguarding data and privacy in intelligent transportation applications and emphasizes the potential of cloud-based solutions to resolve these issues. Organizations can protect sensitive data and user privacy by employing encryption, access controls, threat detection mechanisms, and privacy protection measures. Adopting these cloud-based solutions will encourage the extensive adoption of intelligent transportation applications while infusing users and stakeholders with confidence.

**Key words:** Data security, Privacy protection, Intelligent transportation applications, Cloud-based solutions, Confidentiality, Integrity, Access control.

**1. Introduction.** Intelligent transportation systems have transformed our mobility and interaction with transportation networks. By utilizing sophisticated technologies such as cloud computation, these systems provide numerous benefits, including improved traffic management, increased safety, and enhanced efficiency [11]. However, the growing reliance on cloud-based solutions raises security and privacy concerns. This article will discuss the challenges of safeguarding data and privacy in intelligent transportation applications and how cloud-based solutions can address these issues [23].

Intelligent transportation systems have revolutionized how we move and interact with transportation networks. By leveraging advanced technologies such as cloud computing, these systems offer numerous benefits, including improved traffic management, enhanced safety, and increased efficiency [35]. However, the increasing reliance on cloud-based solutions also raises concerns about data security and privacy.

As intelligent transportation applications generate vast amounts of data, ranging from real-time traffic information to personal user details, safeguarding this data is crucial to prevent unauthorized access, data breaches, and potential misuse [6]. Cloud-based solutions provide a robust framework to protect sensitive information by employing various security measures [21].

This article will explore the challenges of protecting data and privacy in intelligent transportation applications and discuss how cloud-based solutions can address these concerns. We will delve into the importance of data security, including encryption and access control mechanisms, secure data storage and backup options, and threat detection and intrusion prevention measures. Additionally, we will examine the significance of privacy protection, including anonymization and pseudonymization techniques, privacy by design principles, and data governance and compliance frameworks. By implementing these cloud-based solutions, organizations can mitigate risks, enhance user trust, and foster the widespread adoption of intelligent transportation applications. It is imperative to balance the advantages of cloud computing and the need for robust data security and privacy protection.

This article explores the significance of data security and privacy in intelligent transportation applications, investigates how cloud-based solutions can effectively address these concerns, and highlights the role of encryption, access controls, threat detection, privacy-by-design principles, and data governance frameworks in ensuring a secure and privacy-aware environment. The article focuses on applying various data security and privacy measures in the context of intelligent transportation systems, mainly through cloud-based solutions. While the provided abstract and content outline provides a comprehensive overview of the topic, they lack specific mathematical or analytical components. However, it's possible to incorporate mathematical concepts or analyses in certain areas. The article mentions robust encryption techniques like Advanced Encryption Standards (AES). A potential mathematical analysis could involve explaining the mathematical basis of AES

encryption, including the substitution-permutation network (SPN) structure and the mathematical operations (such as substitution and permutation) involved in the encryption process. The article briefly discusses real-time threat detection using machine learning algorithms. A mathematical analysis could delve into the types of machine learning algorithms used, such as anomaly detection algorithms (like Isolation Forest) and their mathematical foundations, including the formulation of anomaly scores and decision boundaries. The article mentions Role-Based Access Control (RBAC) and multi-factor authentication. A mathematical analysis might involve explaining the mathematical principles behind RBAC, such as role hierarchies, permissions matrices, and the mathematical representation of access control policies. The article introduces privacy-enhancing techniques like data anonymization and pseudonymization. A mathematical analysis could explain the concepts of k-anonymity, l-diversity, and differential privacy, showcasing the mathematical mechanisms to protect individual privacy while retaining data utility. The article discusses integrating privacy into system design. A mathematical analysis could include formal methods for specifying and verifying privacy properties in system architectures, ensuring that privacy guarantees are mathematically upheld throughout the system's lifecycle. The article mentions data governance frameworks. A mathematical analysis could involve discussing data classification methods, retention policies, and the mathematical modelling of data lifecycle management processes.

**2. Statement of Problem.** Intelligent transport applications play a crucial role in current transport systems, generating and processing immense quantities of sensitive data. Nevertheless, ensuring the security and confidentiality of this data presents significant challenges. Unauthorized access, data breaches, and potential misconduct can have severe repercussions, eroding user confidence and system integrity. To resolve these issues, it is necessary to investigate how cloud-based solutions can provide a secure framework for protecting sensitive data in intelligent transportation applications.

This article examines the significance of data security in intelligent transportation applications and the mechanisms and solutions provided by cloud platforms to ensure confidentiality, integrity, controlled access, and regulatory conformance. This article endeavours to contribute to developing strategies and best practices for implementing secure and privacy-enhancing cloud-based solutions in intelligent transportation applications by delving into the topics mentioned above. The article will specifically address the following research questions:

- Given the nature of the data involved and the potential threats, what is the significance of data security in intelligent transportation applications?
- How can cloud-based solutions address intelligent transportation applications' data security and privacy concerns?
- In the context of intelligent transportation applications, what mechanisms and solutions do cloud platforms provide to assure confidentiality, integrity, controlled access, and regulatory compliance?

This article examines the challenges and opportunities associated with data security and privacy in intelligent transportation applications by reviewing these research questions. The findings will contribute to advancing this field's knowledge and provide industry professionals, policymakers, and researchers with practical guidance for instituting effective security measures—privacy-enhancing practices in cloud-based intelligent transportation systems.

**3. Research Objectives.** To demonstrate how cloud-based solutions provide a robust framework for safeguarding sensitive data, assuring confidentiality, integrity, controlled access, and regulatory compliance in intelligent transportation applications. To demonstrate the importance of data security and privacy protection in cloud-based intelligent transportation applications and to emphasize the solutions and mechanisms offered by cloud platforms to address these issues effectively.

**4. Research Methodology.** A systematic literature review methodology will investigate the importance of data security in intelligent transportation applications and the role of cloud-based solutions in addressing data security and privacy concerns. This research method involves a systematic and rigorous approach to identifying, selecting, and analysing relevant scholarly articles, research papers, and publications in the field.

**Identification of Research Objective and Research Questions.** The problem statement's research objective and questions have been directed to the literature review. Examining the mechanisms and solutions cloud platforms offer to guarantee confidentiality, integrity, and regulated access.

**Search Strategy.** To identify pertinent literature, a thorough search strategy is developed. Multiple electronic databases, including ACM Digital Library, IEEE Xplore, ScienceDirect, and Google Scholar, will be searched using keywords and search terms pertinent to data security, privacy, cloud computing, and intelligent transportation application research. Additional sources, such as conference proceedings, journals, and relevant reports, will also be considered. The following are the search terms used:

- Data security, intelligent transportation applications, privacy protection, and cloud-based solutions.
- Integrity or confidentiality or access control; regulatory compliance and cloud computing.
- Data storage and backup and cloud-based solutions.
- Threat detection or intrusion prevention.
- Anonymity or pseudonymization and privacy by design.
- Frameworks for data governance and compliance, sophisticated transportation applications.
- User confidence or privacy safeguards and cloud-based solutions.

**Inclusion and Exclusion Criteria.** Inclusion and exclusion criteria were formulated to ensure the selection of pertinent studies. These criteria include the publication date within the last two decades, the topic-related research focus, the methodologies utilized, and the applicability of the findings to the research questions addressed. The selection procedure is conducted in an organized and open manner.

**Study Selection.** The identified studies were subjected to a two-step selection process based on the inclusion and exclusion criteria defined. The relevance of titles and abstracts to the research questions is determined initially through a preliminary screening. The full-text articles of potentially relevant studies were evaluated for ultimate inclusion in the literature review.

**Data Extraction and Analysis.** Data extraction required systematically collecting pertinent information from the chosen studies, such as research objectives, methodologies, critical findings, and limitations. The extracted data were organized and synthesized to identify common themes, trends, and patterns associated with data security in intelligent transportation applications and the role of cloud-based solutions.

**Critical Appraisal.** To ensure the findings' validity and dependability, the chosen studies' quality and importance were critically evaluated. This evaluation assessed the research design, methodology, data analysis techniques, and authors' credibility.

**5. Synthesis and Interpretation.** The data synthesized from the selected studies were analyzed, summarized, and interpreted to address the research questions. The findings were presented coherently and logically, highlighting the key insights, gaps, and suggestions. The results were reviewed regarding the study's goals and questions. Research gaps are highlighted, and the literature's limitations are discussed. Finally, the findings and implications of the systematic literature evaluation regarding the value of data security and the function of cloud-based solutions for intelligent transportation applications are summarized. This research provides a comprehensive overview of the existing literature on data security in intelligent transportation applications and the effectiveness of cloud-based solutions in addressing data security and privacy concerns by employing a systematic literature review methodology. The systematic approach assures the review procedure's rigour, transparency, and reproducibility, contributing to the credibility and validity of the research results.

**6. The Importance of Data Security.** Intelligent transport solutions provide substantial quantities of data, including real-time traffic information as well as user-specific data [18]. It is essential to ensure the safeguarding of this data against unauthorised access, breaches, and possible abuse. Cloud-based solutions provide a comprehensive framework for protecting sensitive data via the implementation of diverse security mechanisms.

Data security is of utmost importance in intelligent transportation systems, since it plays a critical role in safeguarding sensitive data [8]. According to [20], these apps produce substantial quantities of data, including real-time traffic updates, vehicle tracking data, user-specific information, and financial transactions. In order to uphold the confidence of users and stakeholders and mitigate the risks of unauthorised access, data breaches, and possible exploitation, it is essential to prioritise the preservation of confidentiality, integrity, and accessibility of this data [13].

Cloud-based solutions provide a rigorous foundation for data protection, which becomes advantageous for intelligent transport applications. Within this particular setting, many elements serve to underscore the significance of data security.

This research underscores the need to ensure the availability and ease of access to data. The evaluation of access control policies is supported by various mathematical foundations. These include formal languages that are utilised to specify policies, algorithms that enforce these policies, system models that verify their compliance with security-related properties, Bayesian probability and statistical analysis, homomorphic encryption techniques that enable computations on encrypted data, Z-score or Mahalanobis distance for measuring statistical significance, partially ordered sets and lattice structures for organising access control relationships, and the mathematical principles underlying encryption algorithms such as Rivest–Shamir–Adleman or Advanced Encryption Standard.

The method is sometimes referred to as Rivest–Shamir–Adleman, after the surnames of its creators. The aforementioned cryptographic technique is extensively used in asymmetric encryption for the purpose of ensuring the safe transmission of data and the generation of digital signatures. The RSA encryption algorithm relies on the mathematical features inherent in big prime numbers and the inherent challenge of factoring them. The technique produces a pair of cryptographic keys, consisting of a public key and a private key. The public key is used for the purpose of encryption, while the private key is utilised for decryption.

The Advanced Encryption Standard (AES) is a widely used symmetric encryption technique that has emerged as the prevailing standard for securing data via encryption. The Advanced Encryption Standard (AES) is a cryptographic algorithm that acts on blocks of data. It provides support for key lengths of 128, 192, or 256 bits. The technique uses a range of mathematical processes to provide strong data encryption, including substitution-permutation networks (substitutions and permutations) and critical expansion.

The RSA and AES algorithms are well-recognised cryptographic techniques that play a crucial role in safeguarding data confidentiality and security across a range of applications, such as communication, data storage, and digital signatures.

**6.1. Confidentiality.** Intelligent transport systems often manage very sensitive data, including user identities, payment details, and location information. The encryption technologies used in cloud computing effectively obfuscate the data, making it incomprehensible to those who lack proper authorization. The study conducted by [14]. Encryption ensures that data remains safe and unavailable to unauthorised parties, even in the event of interception, unless the appropriate decryption keys are used. Cloud-based solutions have many benefits in intelligent transportation systems, including enhanced scalability, real-time data processing capabilities, and cost-effectiveness. Nevertheless, the growing dependence on cloud services presents a significant obstacle in safeguarding sensitive data and maintaining confidentiality. The incorporation of cloud-based technologies in intelligent transportation necessitates the implementation of strong safeguards to ensure the protection of privacy and security for both users and the transportation infrastructure. In the subsequent discussion, we will examine the significance of maintaining confidentiality in cloud-based intelligent transportation systems and investigate essential approaches for safeguarding sensitive data. Transportation apps that are considered innovative often have the need to regularly gather and retain personal information that is of a sensitive nature. This information may include GPS coordinates, details pertaining to vehicle identification, and user preferences. Ensuring confidentiality is of utmost importance in order to mitigate unauthorised access and improper use of this data, hence safeguarding the privacy rights of persons.

According to [36], cloud-based solutions play a crucial role in the administration and control of traffic. Preserving the confidentiality of traffic data, specifically pertaining to real-time vehicle flow and congestion information, is of utmost importance in order to mitigate any malevolent actions that may impede the transportation network or compromise public safety. Collaboration between transport authorities and research institutes may be facilitated via the implementation of novel projects inside cloud platforms. The use of confidentiality measures is necessary in order to mitigate the risk of unauthorised dissemination of intellectual property, proprietary algorithms, and experimental data. The use of effective encryption algorithms for data at rest and in transit serves to guarantee the data's unreadability in the case of unauthorised access. The encryption protocols TLS and AES are extensively used in many applications. By implementing robust access control measures and using multi-factor authentication, the organisation may ensure that only those with proper

authorization are granted access to sensitive data and essential system components. The implementation of role-based access control (RBAC) is a method that limits user access by considering their designated tasks and obligations [17]. This approach mitigates the risk of unauthorised individuals gaining access to sensitive data. APIs play a crucial role in the integration of various components within intelligent transportation applications. The use of authentication and access restrictions in securing APIs serves the purpose of mitigating unauthorised access to confidential backend data.

The implementation of a data minimization plan involves the collection and retention of just the necessary data for specific purposes. The act of limiting the exposure of sensitive information and reducing the possible consequences of a data breach is a recommended practice. The implementation of continuous monitoring and frequent audits of cloud systems facilitates the timely detection of potential vulnerabilities and unauthorised activities. Real-time alerts may be configured to identify and flag any atypical conduct. The implementation of robust data deletion measures serves to mitigate the risks associated with data leakage and illegal access to outdated information. The preservation of confidentiality is of paramount significance due to the sensitive nature of the data used in cloud-based intelligent transportation systems, encompassing personal information, traffic management particulars, and private research data. Transportation authorities and cloud service providers can establish user confidence, cultivate public trust, and effectively implement intelligent transportation systems without compromising data privacy and security by implementing a comprehensive approach to safeguard confidentiality. This approach encompasses various measures such as data encryption, robust access controls, and routine monitoring.

The use of mathematical models holds considerable importance in elucidating the underlying principles of encryption algorithms. These models aid in comprehending many mathematical concepts, including modular arithmetic and exponentiation employed in RSA encryption, as well as the substitution and permutation operations involved in AES encryption. This academic discussion pertains to the mathematical principles behind the production of cryptographic keys. It encompasses several themes such as prime number generation, modular inverse computation, and the mathematical foundation of crucial pairings in the context of asymmetric encryption. The encryption mechanism known as Transport Layer Security (TLS) is generally acknowledged in academic literature as being extensively used. The mathematical foundations of Transport Layer Security (TLS) protocols include fundamental principles such as asymmetric key exchange, including the Diffie-Hellman algorithm, as well as symmetric encryption techniques to ensure the safe transfer of data. The mathematical principles behind Role-Based Access Control (RBAC) include the formal depiction of roles, permissions, and user assignments via the use of set theory or graph theory. Various statistical techniques may be used for the purpose of anomaly identification. These strategies include the establishment of thresholds using measures like as mean and standard deviation, as well as more sophisticated approaches like clustering and outlier detection. This study focuses on the development of mathematical models that aim to quantify the sensitivity of data and determine the least amount of data necessary for specific tasks. These models take into account several elements, such as entropy and information theory, in order to accurately assess the level of sensitivity and the amount of data needed. The use of time series or pattern recognition techniques in mathematical analysis includes the utilisation of statistical approaches to analyse logs and identify trends that might potentially signify security breaches. Cryptographic methods used for ensuring safe data deletion include the utilisation of algorithms such as the Advanced Encryption Standard (AES) in a designated mode, such as AES-CTR, to facilitate the process of data wiping.

**6.2. Integrity.** The preservation of data integrity is essential to ensuring that data stays unaltered and free from corruption over its entire lifespan. Cloud systems use checksums and digital signatures as mechanisms to safeguard the integrity of data during its storage or transfer processes [22]. The utilisation of mathematical models holds considerable importance in determining the characteristics of secure hash functions, including collision resistance and preimage resistance. These properties guarantee that the task of discovering two distinct inputs that yield identical hash values or reconstructing the initial input from its hash, is computationally impractical. The cryptographic hash functions include mathematical characteristics, namely the avalanche effect and the challenge of discovering collisions, which guarantee that even little changes in the input data provide hash values that are substantially distinct. The mathematical principles behind MAC algorithms, such as HMAC (Hash-based Message Authentication Code), include the integration of hash functions and secret

keys in order to guarantee the integrity of data and mitigate the risk of unauthorised modifications. The mathematical underpinnings of Public fundamental Infrastructure (PKI) include several fundamental components, such as digital certificates, public-private key pairs, and the mathematical algorithms used in the generation and authentication of digital signatures. The mathematical concepts that form the foundation of the security of cryptographic methods. For example, one may elucidate the challenges associated with the factorization of large semiprime integers in the context of the RSA encryption scheme, as well as the discrete logarithm issue as it pertains to specific cryptographic protocols such as Diffie-Hellman. Mathematical analysis offers the potential to investigate many strategies, such as Merkle Trees, that use hash functions to effectively check the consistency and integrity of extensive data structures. The use of mathematical techniques for the identification and rectification of faults in sent data plays a crucial role in preserving the integrity of data in the presence of transmission defects. The birthday paradox is a mathematical notion that has implications for cryptography techniques. This paradox serves as an illustration of the likelihood of two distinct inputs yielding the same hash result, hence emphasising the need for robust cryptographic hash algorithms to ensure data integrity. The inclusion of mathematical principles and analytical approaches might facilitate a more profound comprehension of the methods by which integrity is preserved in cloud-based systems, using several cryptographic methodologies and data verification processes.

**6.3. Access Control.** The management of data access has significant importance in the context of intelligent transportation systems. Cloud-based solutions enable organisations to effectively deploy strong access control techniques, including multi-factor authentication and role-based access restrictions. The use of these techniques serves to provide stringent controls over data access and modification, hence mitigating the potential for unauthorised individuals to gain in or manipulate the data [25]. A comprehensive examination of the subject matter may include elucidating the mathematical underpinnings of set theory, which serves as the fundamental framework for Role-Based Access Control (RBAC). The ideas of sets, subsets, intersections, and unions are foundational principles in set theory that correspond to the allocation of user roles and permissions. The use of graph theory may be essential in the development of access control systems. A potential avenue for mathematical investigation is the examination of access control via the use of graph theory. In this context, vertices would symbolise users, resources, and permissions, while edges would denote the connections and associations between them. Access control often entails the process of making determinations predicated upon specific criteria. Boolean algebra, a mathematical discipline concerned with binary variables and logical processes, is pertinent in this context. Boolean expressions have the potential to elucidate access control rules and situations via the use of mathematical analysis. Access control techniques often include discrete decision-making and logical activities. The use of discrete mathematics principles, such as permutations and combinations, enables the representation of various access situations and facilitates the examination of potential combinations of permissions and roles.

Formal methods include a set of mathematically rigorous approaches that are used to describe, verify, and validate software and systems. A mathematical analysis may be used to examine the utilisation of formal techniques in order to quantitatively demonstrate the satisfaction of specified security features by access control rules, hence assuring their accurate implementation. The article discusses the concept of multi-factor authentication. A mathematical examination may explore the mathematical principles behind cryptographic methods used in multi-factor authentication (MFA), including the utilisation of cryptographic hash functions to create one-time passwords and the mathematical characteristics associated with public-private key pairs. The process of access control often includes the development and administration of cryptographic keys. The security of several cryptographic methods used in access control relies on fundamental notions in number theory, including prime numbers, modular arithmetic, and the discrete logarithm issue. The assessment of access control strategies may include the use of probabilistic models. A mathematical analysis may be used to examine the application of probability theory in evaluating the probability of certain access situations or assessing the possible consequences of policy alterations. Including these mathematical ideas and analyses may contribute to a more holistic understanding of the technological underpinnings of access control methods in cloud-based intelligent transportation systems. .

**6.4. Threat Detection and Intrusion Prevention.** Cloud systems use advanced security methods to identify and mitigate possible security risks. In the realm of cloud infrastructure, the continuous monitoring of

network traffic and activity is facilitated by intrusion detection systems, firewalls, and anomaly detection algorithms. According to [28], the occurrence of any potentially suspicious behaviour triggers an alarm, prompting the implementation of appropriate countermeasures in order to mitigate the risk of unauthorised access and data breaches. The article discusses the use of anomaly detection methods. A mathematical study may include an exploration of statistical concepts such as the mean, standard deviation, and Gaussian distributions, which are often used to characterise typical patterns of activity. Alerts for suspicious behaviour may be triggered by deviations from these patterns. The process of intrusion detection often incorporates machine learning methodologies. Machine learning methods, such as decision trees, support vector machines, and neural networks, may be elucidated via mathematical analysis. These algorithms are designed to discern patterns that signify infiltration, hence enhancing the security of systems. The process of detecting anomalies often involves the analysis of time series data. A mathematical examination may be conducted to investigate signal processing methodologies for the purpose of preprocessing and analysing time-dependent data, including but not limited to Fourier transforms and wavelet analysis. Bayesian networks can effectively represent intricate interconnections between variables. A mathematical examination may be conducted to explore the mathematical aspects of Bayesian networks, which serve as a means to express interdependencies among various events and facilitate probabilistic intrusion detection. The use of information theory principles, such as entropy and mutual information, enables the quantification of the level of randomness or predictability inherent in data. A mathematical study has the potential to explore the use of these notions in detecting deviations from anticipated patterns in network traffic or system behaviour. The article discusses the monitoring of traffic inside cloud infrastructure. The field of graph theory is capable of representing and analysing various network architectures. A mathematical examination might be conducted to investigate the use of graph theory principles, such as nodes, edges, and connectedness, in the field of network analysis. This study would aim to uncover atypical patterns or nodes exhibiting aberrant behaviour. The article discusses the implementation of alert systems that are triggered by the detection of potentially suspicious activities. Bayesian inference is a statistical method that entails the revision of probability in light of new data. A mathematical study may be conducted to explore the use of Bayesian inference in updating the probability of an event being classified as an incursion, as further data is acquired. An inquiry might be conducted to examine the use of queueing theory in the context of intrusion prevention. Queueing models are valuable tools for the prediction and analysis of system behaviour under varying traffic loads. They may assist in the identification of atypical patterns that may signify an intrusion attempt. The identification of potential risks often entails the discernment of recurring behavioural patterns. A mathematical analysis may include the exploration of pattern recognition methodologies, such as clustering algorithms or hidden Markov models, which are used to detect repetitive patterns within datasets. The use of mathematical ideas and analysis may provide valuable insights into the technological components of threat detection and intrusion prevention systems within cloud-based intelligent transportation applications.

**6.5. Data Storage and Backup.** Intelligent transportation applications benefit from cloud-based solutions' scalable and dependable data storage options. Data can be stored on redundant and geographically dispersed servers, which reduces the risk of data loss caused by hardware malfunctions or natural disasters. Regular data backups enhance data resiliency and facilitate rapid recovery during unanticipated events [16]. The article discusses the concept of redundant servers. A mathematical analysis has the potential to investigate several ideas derived from probability and statistics in order to evaluate the probability of failure for specific components as well as the whole system. The quantification of concepts such as mean time between failures (MTBF) and mean time to repair (MTTR) allows for the mathematical design of systems that possess desirable degrees of fault tolerance. The article discusses the use of servers that are geographically scattered.

A mathematical study may include the exploration of principles derived from graph theory or geometry in order to optimise the distribution of data among servers, hence achieving efficient data access, minimising latency, and enhancing fault tolerance. The implementation of redundancy in storage often incorporates strategies such as erasure coding. The mathematics behind Reed-Solomon codes, which are used in data storage systems for error detection and repair, may be elucidated by a rigorous mathematical analysis. This analysis serves to guarantee the integrity of data, even in the event of server failures. The article discusses the need to implement frequent data backups. Probability theory may be examined via a mathematical study within the framework of data loss and backup techniques. This may include the computation of the chance of data loss

as time progresses, as well as the identification of the most advantageous backup frequency and redundancy measures. Queueing theory may be used to analyse the optimal utilisation of resources, including bandwidth and storage capacity. A mathematical study may be used to investigate the utilisation of queueing models in the allocation of resources for data storage and retrieval. This analysis takes into account many elements such as data access patterns and system loads. Although not expressly stated in the text, it is worth noting that data deduplication is a widely used approach in the field of data storage. A mathematical study may be conducted to explore the mathematical concepts behind data deduplication techniques, which aim to detect and remove redundant data in order to enhance storage efficiency.

Data compression is often used in order to achieve efficient storage. A mathematical study may delve into the underlying principles of data compression methods, such as Huffman or arithmetic coding, by examining information theory topics like entropy and coding theory. Mathematical models are used for the evaluation of the dependability and accessibility of storage systems. For instance, the use of Markov models or Petri nets enables the modelling of storage components' behaviour and the anticipation of system dependability across various circumstances. The article discusses the phenomenon of rapid recuperation in the face of unforeseen circumstances. An investigation in mathematics may be conducted to examine the correlation between the frequency of backups, recovery time goals (RTO), and the possible loss of data (recovery point objective - RPO), therefore offering valuable insights for the development of efficient backup systems. The use of mathematical ideas and analyses may enhance comprehension of the technological concerns involved in guaranteeing dependable and robust data storage and backup procedures inside cloud-based intelligent transportation systems.

**6.6. Compliance with Regulations.** Numerous nations have stringent data security and privacy regulations, such as the European Union's General Data Protection Regulation (GDPR). Cloud platforms provide organizations with frameworks and instruments for maintaining regulatory conformance. These tools assist in defining data retention policies, consent management mechanisms, and audit trails for data access, ensuring compliance with applicable regulations [15]. In intelligent transportation applications, data security is of the utmost importance. Cloud-based solutions provide a robust framework for safeguarding sensitive data, ensuring its confidentiality, integrity, and controlled access. Organizations can mitigate risks and preserve users' and stakeholders' confidence in intelligent transportation systems by instituting encryption, access controls, threat detection mechanisms, and robust storage and backup solutions [12]. Mathematical models have the potential to be formulated in order to guarantee adherence to legislation, such as the General Data Protection Regulation (GDPR). These models may include the establishment of mathematical principles and criteria that govern data processing operations in order to ensure compliance with legal obligations. Privacy impact assessments include the evaluation of possible privacy hazards that may arise from the processing of data. Mathematical analysis may be used to estimate privacy concerns by using probability and effect evaluations. This approach aids organisations in making well-informed choices about their data handling practices. The article discusses the implementation of audit trails as a means of monitoring and documenting data access activities. A mathematical study may be used to investigate graph theory principles in order to construct a model for data provenance, which facilitates the tracing and documentation of data lineage and historical information, hence enabling traceability and accountability. Temporal logic is a valuable tool for the expression and verification of material features, including but not limited to data retention regulations and permission expiry. A mathematical study may be conducted to examine the use of temporal logic in formalising and verifying compliance rules across a period of time. Consent management encompasses the practice of monitoring and recording user consent pertaining to the processing of their data. A mathematical study may be used to investigate the utilisation of graph theory in representing the interconnections among users, data processing activities, and permission status, with the aim of facilitating effective consent management. Formal verification procedures include the use of mathematical methodologies to rigorously ascertain the adherence of systems to defined attributes. A mathematical study may be used to examine how conventional practices might effectively exhibit adherence to regulatory mandates, therefore guaranteeing that data processing operations conform to legal benchmarks. The essay discusses the importance of reducing risks and maintaining trust. Probability theory may be used in a mathematical study to examine risk assessment, specifically in the calculation of the probability of security breaches or non-compliance incidents. This analysis can provide valuable insights for developing effective risk management methods. Privacy metrics may be measured in order to evaluate the level of privacy safeguard-



ing in the context of data processing. A mathematical examination may explore the computation and use of privacy metrics, such as k-anonymity or differential privacy parameters, in order to guarantee adherence to privacy standards. Organisations often encounter the need for data sharing while maintaining compliance with regulations. The use of game theory principles presents a viable approach for the modelling of interactions among entities engaged in data sharing. This entails a comprehensive analysis of the incentives and legal limitations that influence the behaviour of these parties. The essay discusses the importance of adhering to regulatory standards. The use of Bayesian inference allows for the continual monitoring and updating of compliance evaluations, taking into account fresh evidence and adjusting to changes in data processing operations and regulatory requirements. The inclusion of these mathematical principles and analyses may provide a more holistic understanding of the technological and legal dimensions involved in guaranteeing regulatory adherence and safeguarding data integrity within cloud-based intelligent transportation systems.

**7. Encryption and Access Control.** Encryption, which entails encapsulating data to make it opaque to unauthorized users, is crucial to data security. Cloud platforms provide robust encryption mechanisms to safeguard data during transmission and storage [1]. Furthermore, the implementation of access control techniques, such as multi-factor authentication and role-based access restrictions, may effectively limit access to those who have been granted authorization. The mathematical underpinnings of cryptography include several principles, including encryption algorithms, cryptographic keys, and the intricacies of the encryption and decryption procedures. This includes mathematical techniques, such as modular arithmetic, which are used in encryption methods. The article discusses the implementation of strong encryption techniques. The administration of cryptographic keys plays a significant role in the field of encryption. A mathematical analysis may be used to investigate topics like key generation, key exchange protocols (e.g., Diffie-Hellman), and the mathematical principles that guarantee the security of cryptographic keys. The notion of Public Key Infrastructure (PKI) encompasses mathematical principles associated with asymmetric cryptography, including the mathematical correlation between public and private keys, as well as the computational intricacy of certain mathematical problems that form the foundation of PKI's security measures. The discussion might revolve around several notions in information theory, including entropy and the notion of complete secrecy, as first proposed by Claude Shannon. These principles elucidate the theoretical boundaries of safe communication and the inherent indeterminacy that encryption cannot fully eradicate. The field of number theory plays a crucial role in the development and implementation of the RSA encryption algorithm. The study of encryption may include an examination of the mathematical aspects pertaining to RSA encryption, including the mathematical characteristics of prime numbers, modular exponentiation, and the challenges associated with the factorization of large semiprime integers. Elliptic Curve Cryptography (ECC) is a cryptographic method that falls under asymmetric encryption. An examination of the algebraic structures of elliptic curves and their applications in cryptographic operations may be undertaken via mathematical study. Although not expressly stated, steganography is a technique that entails concealing the existence of information by embedding it into other data. The importance of information theory and statistical analysis in constructing efficient steganographic methods may be investigated via a mathematical study. Access control techniques often include logical processes. The field of Boolean algebra, which pertains to the manipulation of binary variables and logical processes, is applicable in this context. A mathematical analysis may be used to elucidate the manner in which Boolean expressions can establish access control rules and criteria. Graph theory may be used to depict access control schemes. A mathematical study may be conducted to examine the use of graph theory principles, such as nodes and edges, in the modelling of access connections and permissions. The use of finite automata theory extends to the modelling of access control policies and the establishment of the underlying logic for implementing certain licences and limitations. Formal methods include a set of mathematically rigorous approaches used for the purpose of system verification. A mathematical analysis may be used to examine the applicability of conventional approaches in validating the accuracy of access control rules and procedures. The inclusion of mathematical ideas and research may contribute to a more profound comprehension of the technological foundations of encryption, access control, and data security procedures inside cloud-based intelligent transportation systems.

**7.1. Secure Data Storage and Backup.** Cloud-based solutions provide intelligent transportation applications with reliable and scalable data storage alternatives. According to [31], storing data on redundant and geographically distributed servers may mitigate the potential risks of data loss resulting from hardware

breakdowns or natural catastrophes. The use of regular data backups serves to improve the resilience of data and expedite the process of recovery in unforeseen circumstances. Mathematical models can be constructed to evaluate the dependability and accessibility of data storage systems. Dependability engineering principles, such as failure rates, may be employed to quantitatively assess the trustworthiness of individual servers and the overall system architecture. The use of queueing theory is relevant in the modelling of storage systems since it enables the analysis of many performance metrics, such as waiting times for data access, system utilisation, and reaction times under varied workloads. The article discusses the potential for data loss from hardware problems or natural calamities. Probability theory enables the evaluation of the probability of these occurrences transpiring and the computation of their possible consequences on data availability. The topic of regular data backups is being addressed. Mathematical analysis encompasses the computation of recovery time goals (RTO) and recovery point objectives (RPO) by leveraging backup frequency, hence facilitating the optimisation of backup techniques. Erasure coding encompasses the use of data redundancy methodologies. A mathematical study may be conducted to explore the mathematical principles behind erasure coding techniques, which generate duplicate data pieces to facilitate data recovery in the event of partial data loss. The presence of geographically scattered servers suggests the use of data replication. Graph theory may be employed to represent and analyse data replication schemes, guaranteeing the optimal dissemination of data copies over several servers to minimise latency. Bayesian networks have the capability to effectively define and analyse interdependencies and potential hazards within the context of disaster recovery planning. A mathematical study may investigate the use of Bayesian networks in evaluating the probable consequences of different catastrophe scenarios. Data deduplication is a storage optimisation technique that entails identifying and eliminating redundant data. Using information theory principles, such as entropy, allows for quantifying data redundancy and optimising deduplication algorithms. The article discusses the use of servers that are distributed across different geographical locations. Geometric principles, such as distance metrics, may be used to optimise the positioning of servers to achieve data redundancy and mitigate the impact of geographical disturbances. Graph colouring methods distribute data to servers so duplicate copies are kept on distinct servers, reducing the likelihood of simultaneous loss. Markov models enable the examination of server behaviour by assessing the possibilities of failure and recovery. This analytical approach facilitates predicting system performance and behaviour over a certain period. Using mathematical ideas and analyses may provide valuable insights into the technological issues involved in developing dependable and robust data storage and disaster recovery methods within cloud-based intelligent transportation systems.

**7.2. Threat Detection and Intrusion Prevention.** Cloud platforms utilize advanced security measures to detect and mitigate potential hazards. Intrusion detection systems, firewalls, and anomaly detection algorithms constantly monitor cloud infrastructure traffic and activity [27]. Any potentially illicit behaviour triggers an alarm system, prompting the implementation of appropriate countermeasures to thwart unauthorised entry and safeguard against the compromise of sensitive information. The topic of anomaly detection techniques is discussed. Statistical principles, such as the calculation of mean, variance, and standard deviation, may be used to establish benchmarks for typical behavioural patterns and detect departures from these patterns, which may indicate possible risks. Machine learning methodologies, such as clustering algorithms or support vector machines, may be effectively used to identify discernible patterns of behaviour that indicate the presence of intrusion. The mathematical principles underlying the training and use of machine learning models are pertinent in this context. Probability theory, specifically Bayesian inference, may be used to evaluate the likelihood of certain acts being malevolent by integrating past information with newly acquired evidence. Information theory concepts, such as entropy, can quantitatively measure the level of randomness or unpredictability in data. Anomalies often have elevated entropy levels, a characteristic that may be recognised by mathematical means. The topic of discussion pertains to the surveillance of network traffic inside cloud infrastructure. The field of graph theory is used to represent and analyse network topologies. This framework examines many elements, such as nodes, edges, and connectedness, to uncover atypical patterns or nodes exhibiting aberrant behaviour. The anomaly identification process often entails examining and analysing time series data. Time series analysis is a statistical technique encompassing many methods, such as autoregressive integrated moving average (ARIMA) models and exponential smoothing. These methods detect and analyse trends, patterns, and anomalies within organised chronological data. Bayesian networks can effectively represent and capture

the interdependencies between various occurrences. A mathematical study may be conducted to investigate the capacity of Bayesian networks to represent the interconnections among different system activities and their potential as indicators of threats. The identification of concealed data, such as malicious software, may be addressed via steganalysis methodologies. The process entails using mathematical analytic techniques to identify hidden patterns or variations within data, which serve as indicators for potential risks. The field of queueing theory enables the analysis of traffic patterns to detect abnormalities. Mathematical techniques may identify anomalies by comparing observed queue lengths and waiting times with their corresponding predicted values. Mathematical optimisation approaches may be used to establish rules for detecting risks. The regulations mentioned above possess the capability to develop thresholds for identifying suspicious activity by using mathematical analysis of past data. The article discusses the concept of intrusion prevention. Graph colouring algorithms can analyse access control links and detect probable patterns of unauthorised access. Including these mathematical principles and analyses may facilitate a more profound comprehension of the technological underpinnings of threat detection and intrusion prevention techniques in cloud-based intelligent transportation systems.

**7.3. Privacy Protection.** Data security and user privacy protection are of utmost importance in intelligent transport systems. Cloud-based solutions provide a wide range of methods aimed at safeguarding privacy. Differential privacy is a mathematical paradigm that quantifies the extent to which an individual's privacy is compromised when their data is included in a dataset. The use of differential privacy principles may effectively safeguard the privacy of a dataset by minimising the impact of every data point on the overall privacy of the information. The privacy models discussed below pertain to the alteration of data so that it becomes indiscernible within a collective of persons. Mathematical evaluations may be used to investigate the efficacy of various models in maintaining privacy while enabling data usefulness. Using concepts derived from information theory, such as entropy, allows quantifying the degree of information or privacy that is compromised throughout the data exchange process. The examination of information loss may assist in achieving a harmonious equilibrium between data usefulness and the protection of privacy. Mathematical methodologies, such as safe multi-party computing and homomorphic encryption, can facilitate data mining on encrypted data while guaranteeing the preservation of confidentiality for sensitive information throughout the analysis process. Zero-knowledge proofs and rapid multiparty computing are cryptographic protocols designed to maintain data confidentiality by allowing it to be utilised for specific purposes without disclosing the actual contents. The mathematical underpinnings of these methods may be investigated. The Laplace mechanism is used to introduce random perturbations into query replies, safeguarding the anonymity of individual users. The mathematical studies may be directed towards determining the ideal quantity of noise to be added, while simultaneously ensuring the preservation of data usefulness. The introduction of controlled noise derived from entropy may be used to disturb data and provide privacy protection. Mathematical analysis may be used to investigate the optimal level of entropy that can be introduced to strike a balance between preserving privacy and enabling meaningful analysis. These methods include modifying data to eliminate personally identifying information. Mathematical computations may evaluate the efficacy of anonymization techniques in mitigating the potential for re-identification. Secure multi-party computation protocols refer to a computational framework whereby numerous parties collaborate to perform a function while ensuring the confidentiality of their respective inputs. The mathematical principles behind these protocols may be examined to understand their potential use in privacy-preserving analytics. The safeguarding of privacy requires a comprehensive comprehension of the movement of data. Graph theory ideas may be used to represent and analyse data flows, hence enabling the identification of possible privacy risks and the proposal of strategies to strengthen security. Methods such as  $\epsilon$ -differential privacy include monitoring and allocating a privacy budget. Mathematical studies may be used to investigate the optimal allocation of funding to ensure both effectiveness and the preservation of a satisfactory degree of privacy. By integrating mathematical principles and conducting empirical investigations, one may get valuable insights into the technological underpinnings of privacy protection measures used in cloud-based intelligent transportation systems.

**7.4. Anonymisation and Pseudonymisation.** Prior to being stored or processed in cloud-based systems, personal data has the potential to undergo anonymization or pseudonymization techniques to safeguard the identities of users. This methodology entails the replacement of personally identifiable information with distinct identifiers or the deliberate obscuring of the data, so impeding the ability to establish a direct con-

nection between the data and particular people. Entropy from information theory, specifically the Entropy and De-identification framework, may be used to evaluate the extent of information loss that occurs through anonymization or pseudonymization. Mathematical analysis may facilitate the identification of an appropriate equilibrium between the value of data and the protection of privacy. The privacy models of K-anonymity and L-diversity aim to guarantee the indistinguishability of data inside a given group. Mathematical analysis may be used to investigate the influence of these models on the quality of data and the preservation of privacy. Probability theory enables the evaluation of the potential for re-identification by analysing the retained information in anonymized or pseudonymized datasets. Mathematical analysis can calculate the probability of a successful re-identification assault. Through mathematical principles about utility, it becomes possible to examine the extent of information loss that occurs during the process of anonymization or pseudonymization, as well as to assess if the resulting data retains an adequate level of value for analytical purposes. Using graph theory ideas enables the analysis of linkages between anonymized data and possible external data sources, hence facilitating the evaluation of data linkage and re-identification risk. Hash functions can be used for pseudonymizing data. Mathematical analysis enables examining the characteristics of cryptographic hash functions and their implications for safeguarding data integrity and mitigating re-identification vulnerabilities. Mathematical computations may be used to examine the inherent trade-offs between the value of data and the safeguarding of privacy, so aiding in the determination of the optimal degree of anonymization or pseudonymization that simultaneously maximises both goals. Latent Variable Models include the task of identifying latent variables that account for the observable data.

Mathematical analysis may be used to investigate the potential of latent variable models in facilitating pseudonymization processes while maintaining the integrity of data attributes. The use of mutual information, a concept from information theory, allows for quantifying the extent to which pseudonymization reduces the leaking of information. Mathematical methods may be used to quantify the extent to which pseudonymization effectively mitigates information leakage. Statistical Disclosure Control encompasses many techniques aimed at safeguarding against re-identification assaults. Mathematical computations may be used to investigate the impact of noise injection, aggregation, and suppression on the likelihood of re-identification. Mathematical optimisation methods may be used to identify the optimal pseudonymization approach that maximises data value while ensuring robust privacy protection. The integration of mathematical ideas and studies may provide valuable insights into the technological underpinnings of anonymization and pseudonymization processes in intelligent transportation systems hosted on cloud platforms.

**7.5. Privacy by Design.** Incorporating privacy concerns into the design and development of intelligent transportation applications is essential. Cloud-based solutions can enforce privacy by design principles, ensuring privacy safeguards are built into application development [34]. This strategy helps reduce privacy risks and assures compliance with privacy laws. Privacy by Design (PbD) is a proactive and ethical approach that prioritises data protection and privacy throughout the entire lifecycle of intelligent transportation cloud applications [34, 26]. PbD emphasises incorporating privacy concerns into the design, architecture, and development of systems instead of addressing privacy concerns as a supplement [4]. Privacy by Design plays a crucial role in ensuring the confidentiality, integrity, and ethical use of sensitive data in intelligent transportation applications that leverage cloud-based solutions [7]. Privacy by Design stipulates that cloud-based conveyance systems should be designed with data protection and privacy [3]. By being proactive, it is possible to identify and mitigate potential privacy risks and vulnerabilities before they become significant problems. Privacy by Design (PbD) encourages privacy-friendly defaults, assuring that the maximum level of privacy protection is implemented by default to user data. Users should not be required to configure their privacy settings to protect their data.

Privacy by Design emphasises that data protection measures should not hinder cloud-based transit applications. Instead, Privacy by Design (PbD) enhancing technologies [5], Privacy by Design (PbD) incorporated seamlessly to facilitate optimal user experiences [37].

Privacy by Design (PbD) considers data security throughout the entire data lifecycle, including data acquisition, storage, processing, and disposal. It safeguards data at every stage, minimising the possibility of unauthorised access.

The transparency and comprehensibility of privacy rules for intelligent transportation apps are crucial in

enabling users to make informed choices about the sharing and utilisation of their data. The recommended approach is to selectively gather and retain just the necessary data for the specific objective, thereby reducing the likelihood of disclosing confidential information. In order to safeguard the identities of individuals while maintaining the research and analytical value of the data, it is recommended to use methods such as anonymization or pseudonymization. To ensure the security of data, it is essential to use encryption techniques throughout both transmission and storage processes. It is also crucial to establish and enforce strong access restrictions to mitigate the risk of unauthorised access.

It is recommended to do privacy impact assessments to detect and address any privacy concerns and to assure adherence to legislative requirements.

Ensure that users provide unequivocal permission to process their data, and enable them to exercise their rights to access, alter, or delete their data as required.

Incorporating Privacy by Design principles is essential in developing and implementing intelligent transit applications that rely on cloud-based technology. Transportation authorities and cloud service providers can foster user confidence, preserve sensitive data, and comply with ethical and legal data protection standards by incorporating privacy concerns. Privacy by Design grants people more agency in managing their data while guaranteeing that intelligent transport advancements contribute to a digital environment that prioritises security and respects privacy. Privacy Impact Assessments (PIAs) do not possess an intrinsic mathematical nature; they include methodical evaluations of possible privacy hazards. Quantitative analysis encompasses the assessment of the potential consequences arising from a privacy breach, specifically about factors such as the sensitivity of the compromised data, the level of confidence placed by users, and the fines imposed by regulatory bodies. Risk quantification is not only reliant on mathematical methods. The process of risk assessment and quantification may include the assignment of numerical values to prospective privacy threats. This may facilitate prioritising mitigation solutions by considering their possible effect and probability. The usability and user experience analysis is not only based on mathematical principles. In reality, user experience evaluation may include quantitative measures such as user satisfaction ratings, interaction durations, and job completion rates. These metrics serve the purpose of ensuring that privacy-enhancing technologies do not impede the overall user experience. Privacy effect measurements are not exclusively based on mathematical principles. Developing privacy impact metrics may include establishing quantitative indicators of privacy safeguarding, such as quantifying the anonymity attained via anonymization methods. Consent Management and Compliance Metrics involve using quantitative analysis to evaluate the efficacy of user consent methods and the level of adherence to regulatory requirements shown by privacy policies.

Data Minimization Metrics refers to a mathematical approach that enables the measurement of the extent to which data exposure is reduced via data minimization measures. This analysis aids in striking a balance between the usefulness of data and safeguarding privacy. Privacy-preserving analytics refers to data analysis while protecting individuals' privacy. The evaluation of privacy-preserving analytics techniques encompasses more than just mathematical considerations. It entails a comprehensive assessment of their performance, including quantitatively comparing the accuracy and insights obtained from analysing raw data with those acquired from analysing privacy-protected data. The administration of encryption keys encompasses more than just mathematical aspects. It includes implementing safe mathematical concepts, such as crucial creation, distribution, and rotation.

Access control metrics are not only based on mathematical principles. Evaluating the efficacy of access controls may include assessing the frequency of unauthorised access attempts and successful breaches. The rates at which individuals actively consent to a specific action or process. The quantification of the proportion of users who voluntarily participate in data processing activities might provide valuable information about their inclination to contribute data and their level of confidence in the privacy safeguards implemented by the system. The study titled "Quantifying the Effects of Privacy Enhancements on System Performance" presents a mathematical framework that may be used to assess the influence of various privacy-enhancing protocols on the operational efficiency of intelligent transportation systems. This analysis encompasses processing speed, storage demands, and computational overhead. Privacy by design emphasises principles and holistic considerations. At the same time, the practical implementation and evaluation of privacy-preserving mechanisms and user experiences may involve quantitative analysis, metrics, and the application of mathematical concepts to address

specific aspects of privacy protection.

**7.6. Data Governance and Compliance.** Cloud platforms frequently provide data governance and compliance management tools and frameworks. These tools enable organisations to define data retention policies, consent management mechanisms, and data access audit trails to maintain regulatory compliance, such as GDPR or CCPA [30], as cloud-based solutions continue to transform intelligent transportation applications, data governance and compliance become of utmost importance [10]. Data governance incorporates the procedures, policies, and guidelines that prescribe how data is managed, accessed, and protected. In contrast, compliance is the observance of applicable laws, regulations, and industry standards governing data privacy and security. Robust data governance frameworks and compliance measures are necessary to ensure ethical and legal practices in cloud-based intelligent transportation applications. Delineating data ownership and designating data management responsibilities within a cloud-based infrastructure ensures accountability and adequate handling of sensitive data. Implementing measures to maintain data accuracy, consistency, and integrity ensures the dependability of the data used in decision-making and research. Developing data collection, storage, retention, and disposition policies reduces the risk of unauthorised access or data breaches by ensuring that data is handled appropriately throughout its lifecycle. Categorising data according to its level of sensitivity enables the application of appropriate security controls and access restrictions, thereby ensuring that sensitive data is adequately protected. When dealing with personally identifiable information (PII) in intelligent transportation applications, compliance with data privacy laws such as the General Data Protection Regulation (GDPR) in Europe or the California Consumer Privacy Act (CCPA) in the United States is crucial [24]. Effective encryption methods, such as Transport Layer Security (TLS), during data transmission between cloud servers and end users serves as a protective measure against possible eavesdropping and interception of data.

The implementation of optimal security practises may be effectively facilitated by adhering to industry-specific standards, such as ISO 27001 for information security management or the NIST Cybersecurity Framework [32]. Conducting Privacy Impact Assessments (PIAs) aids in the identification and mitigation of privacy risks linked to data processing operations, hence assuring the proactive resolution of privacy problems. The regular implementation of internal and external compliance audits serves to assess the effectiveness of data governance practises and suggest areas for improvement. Establishing formal agreements to delineate the rights, obligations, and security prerequisites regarding data use is essential while engaging in collaborations with third-party businesses or exchanging data across diverse parties. Before sharing or doing analysis, it is common practice to anonymize or pseudonymize sensitive data to protect the identity of individuals and mitigate any privacy hazards.

The implementation of data governance and compliance measures is vital to preserve data integrity, protect users' privacy, and guarantee the ethical utilisation of cloud solutions within intelligent transportation systems. Transportation authorities and cloud service providers can cultivate user trust, stimulate innovation, and establish a secure and sustainable smart transportation ecosystem by implementing robust data governance frameworks, complying with relevant data privacy regulations, and embracing secure data-sharing practices. To fully harness the capabilities of cloud-based solutions within the dynamic transportation industry, it is imperative to prioritise the ethical and legal dimensions of data management, focusing on safeguarding data and privacy. Risk assessment and quantification is a multidimensional process that extends beyond quantitative considerations. In addition to numerical analysis, risk assessment includes the evaluation of the possible consequences and probability associated with non-compliance with data protection standards. Quantitative analysis may be utilised to evaluate regulatory fines' financial and reputational repercussions. The audit trail analytical process involves examining data access and alterations, using mathematical, and analytical techniques to identify patterns of unauthorised access, probable breaches, or irregularities in data use.

The mathematical study known as the Data Classification and Sensitivity study is used to assess the sensitivity of various data kinds and establish suitable security measures by their respective levels of secrecy. The measurement of possible privacy risks found in Privacy Impact Assessments (PIAs) may include quantitative evaluation of the probability and consequences of various hazards. The analysis of data retention periods aims to identify the most efficient duration for storing data, considering both regulatory compliance and storage costs. This process often involves using mathematical models that examine many elements, such as patterns of data consumption, legal obligations, and expenditures associated with data storage. The performance metrics

used in compliance audits assess the effectiveness of these audits in finding gaps and possible breaches. These metrics primarily rely on quantitative measures such as audit coverage, detection rates, and the occurrence of false positives and negatives. The present study focuses on doing a quantitative analysis of consent management. The process of assessing user opt-in rates and monitoring user preferences for data processing includes the quantitative examination of user activity. The evaluation of anonymization methods involves using quantitative measures to determine the protection of identity and preservation of data usefulness. These metrics are not only based on mathematical principles. The Quantitative Assessment of Data Sharing Agreements involves the evaluation of security standards and duties included in formal agreements for sharing data. This assessment may consist of using mathematical techniques to compare different security measures. Compliance Score Metrics provides a metric that measures the degree to which an application conforms to relevant norms and standards. This process may include giving numerical values to various compliance criteria. The topic of concern is the accuracy and consistency of data. Metrics are used to measure the precision and reliability of data during its entire lifespan, and they may impact mathematical analysis and the ability to compare data against predetermined standards.

Although not all components of data governance and compliance are inherently rooted in mathematical principles, the use of quantitative analysis, metrics, and modelling may facilitate the evaluation of various measures' risks, effects, and efficacy. These factors may assist in making well-informed judgements on data management and compliance techniques in cloud-based intelligent transportation apps.

**8. Cloud-Based Solutions For Intelligent Transportation Applications.** Technology development has revolutionised numerous industries, including transportation. Intelligent Transportation Systems (ITS) employ cutting-edge technologies to improve transportation's efficacy, safety, and sustainability [9]. Cloud-based solutions have emerged as a crucial enabler for the implementation and success of these intelligent transportation applications. This article will examine the advantages, difficulties, and prospective applications of cloud-based solutions in intelligent transit. The infrastructure of cloud-based solutions is elastic and can scale resources up or down based on demand. This is especially essential for transportation systems with fluctuating loads, such as traffic management during rush hour or peak season [19].

Platforms in the cloud enable the seamless accumulation, processing, and analysis of enormous quantities of real-time data from various sources, such as GPS, sensors, and cameras. This information can optimise traffic flow, monitor road conditions, and provide travellers with current information [29]. Using a cloud-based strategy, transport authorities can reduce infrastructure expenditures because they no longer need to invest substantially in on-premises hardware and maintenance [33]. In addition, cloud services offer pay-as-you-go pricing models, allowing businesses to pay only for the resources they consume [38]. Cloud solutions facilitate data sharing and collaboration between parties, such as government agencies, transportation providers, and third-party developers. This interoperability encourages the development of innovative applications that can solve complex transportation issues. Cloud platforms enable sophisticated traffic management systems that use real-time data analysis to optimise signal timing, control traffic flow, and reduce congestion. These solutions can result in reduced travel durations and enhanced traffic efficiency overall. In Vehicle-to-Everything (V2X) communication, cloud-based solutions facilitate data transmission between vehicles, infrastructure, and other devices. This technology improves road safety by alerting drivers in real-time to potential road hazards and enhancing their situational awareness.

Public transportation routes, schedules, and fleet management can be optimised using cloud-based analytics. This results in enhanced passenger experiences, decreased operational expenses, and increased public transportation utilisation [2]. Connected parking systems in the cloud can direct drivers to available parking spaces, thereby reducing traffic congestion and emissions caused by drivers browsing for parking. Cloud-based predictive maintenance applications utilise sensor data from vehicles and infrastructure to determine maintenance needs. This proactive strategy improves maintenance planning, decreases idleness, and increases the overall dependability of conveyance systems. Cloud-based solutions require adequate security measures to protect sensitive transportation data from unauthorised access and cyber threats.

As transport systems significantly rely on real-time data, cloud service providers must ensure high availability and redundancy to prevent disruptions. Cloud-based solutions rely on dependable internet connectivity and minimal latency, which can be difficult in remote or underdeveloped areas. When utilising cloud-based services,

transportation authorities must comply with applicable data protection and privacy regulations. The advantages of cloud-based solutions for intelligent transportation applications range from increased scalability and real-time data processing to cost savings and collaboration opportunities. By addressing potential obstacles and employing best practices, transportation authorities can unlock the full potential of cloud-based technologies and promote innovation in the intelligent transportation sector, resulting in a safer, more efficient, and more sustainable transportation network. Scalability Analysis is a mathematical model that can be used to analyze the scalability of cloud-based solutions, considering factors like resource utilization, response times, and costs as the system scales up or down. Resource Allocation Optimization is a mathematical optimization technique that can help determine the optimal allocation of resources within a cloud-based transportation system to ensure efficient data processing and analysis. Real-Time Data Processing is a Mathematical algorithm, and models may be employed to process real-time data efficiently, considering factors like data arrival rates, processing speeds, and latency.

Traffic Flow Optimization is a mathematical simulation or model that can optimize traffic flow based on real-time data, aiming to reduce congestion and travel times. Predictive Analytics is a mathematical technique such as regression analysis or machine learning that can be applied to predict future transportation patterns, helping in proactive decision-making. The article discusses the benefits of cloud-based solutions regarding cost savings and improved efficiency. Mathematical metrics could quantify these improvements in percentages, ratios, or other quantitative measures. Mathematical analysis might be used to assess network latency and its impact on real-time data processing in the context of reliable data transmission. Data Sharing Efficiency is a Mathematical model that could be used to assess the efficiency of data sharing and collaboration among different parties, considering factors like data exchange rates and data integrity. Privacy and Data Protection Metrics are quantitative metrics that could be used to assess the level of privacy protection achieved through cloud-based solutions, considering factors like data anonymization effectiveness and compliance with privacy regulations. Cost-benefit analysis is a mathematical analysis that can be employed to perform a cost-benefit analysis of adopting cloud-based solutions, considering both the costs of implementation and the potential benefits of efficiency, scalability, and user satisfaction. Risk Assessment and Mitigation is a mathematical analysis that might assess potential risks associated with cloud-based solutions, quantifying the likelihood and potential impact of disruptions or security breaches. While the content in the text emphasizes the operational and technological aspects of cloud-based solutions, integrating relevant mathematical analyses or models can enhance the depth and precision of the discussion in areas such as optimization, scalability, efficiency, and risk assessment.

**Enhanced Efficiency.** Existing knowledge of cloud-based solutions highlights their ability to improve the efficiency of intelligent transportation systems by providing real-time data processing, traffic management, and decision-making capabilities.

**Scalability.** Cloud computing allows transportation systems to scale resources based on demand, effectively accommodating fluctuating traffic volumes and user requirements.

**Cost Savings.** Adopting cloud-based solutions can reduce infrastructure and maintenance costs, as organizations can leverage the cloud provider's resources instead of investing in expensive on-premises hardware.

**Data Accessibility.** Cloud-based platforms enable seamless data sharing and collaboration among stakeholders, promoting data-driven decision-making and improved transportation services.

**Innovation and Flexibility.** Cloud solutions foster an environment for innovation and experimentation, allowing developers to create and deploy new intelligent transportation applications quickly.

## **9. Drawbacks of Existing Knowledge on Cloud-based Solutions for Intelligent Transportation.**

**Data Security and Privacy Concerns.** Cloud computing introduces security and privacy risks, as transportation data is stored and processed off-site. Breaches or unauthorized access to cloud-stored data can compromise user privacy and lead to data breaches.

**Reliability and Downtime.** Cloud-based solutions rely on internet connectivity, and downtime or disruptions in internet services can affect the availability and performance of intelligent transportation applications.



**Data Residency and Compliance.** Cloud services might operate across different jurisdictions, leading to challenges in ensuring data residency compliance and adhering to relevant data protection laws.

**Dependency on Cloud Providers.** Organizations relying heavily on cloud-based solutions may face vendor lock-in, making it challenging to switch providers or migrate data if needed.

**Network Latency.** For real-time applications, network latency can affect the responsiveness of cloud-based systems, potentially impacting the overall performance of intelligent transportation applications.

**Complexity of Integration.** Integrating existing transportation infrastructure with cloud-based solutions may be complex and require careful planning to avoid disruptions during the migration process. Existing knowledge of cloud-based solutions for intelligent transportation showcases numerous advantages, such as enhanced efficiency, scalability, cost savings, data accessibility, and opportunities for innovation. However, it acknowledges significant drawbacks, including data security and privacy concerns, reliability issues, compliance challenges, vendor lock-in, network latency, and integration complexities. Addressing these drawbacks is crucial for the responsible and successful adoption of cloud-based solutions, ensuring that data and privacy are protected while reaping the benefits of cloud computing in the intelligent transportation domain.

**10. Case Study.** Cloud-based solutions have revolutionised intelligent transportation, which offers increased efficiency and connectivity. However, the increased use of cloud computing has raised significant data security and privacy concerns. This paper presents a series of case studies and illustrations to illuminate the real-world implications and efficacy of cloud-based security measures in intelligent transportation. These case studies emphasise organisations' challenges and practical implementation strategies, providing valuable insights into data protection in this dynamic domain.

### 10.1. Case Study 1: Citywide Traffic Management System.

**Scenario.** A thriving city implements a municipal cloud-based traffic management system to optimise traffic flow, reduce congestion, and improve public safety.

#### Challenges.

*Data Privacy.* The system accumulates voluminous data from traffic cameras, GPS devices, and connected vehicles, raising privacy and anonymity concerns.

*Data Security.* Protecting real-time traffic data and system integrity against cyber threats and potential hijacking attempts is crucial for preventing disruptions and ensuring public safety.

#### Implementation Strategies.

*Anonymization.* To ensure individual vehicles' and pedestrians' privacy, the city employs sophisticated anonymization techniques that aggregate and conceal personal data to protect user identities.

*Secure Data Transmission.* Using robust encryption protocols for data in transit ensures the confidentiality and integrity of data transmitted between the cloud platform and connected devices.

**Results.** While addressing data privacy concerns, the cloud-based traffic management system optimises traffic flow, reduces congestion, and enhances public safety in general. The anonymization techniques and secure data transmission safeguard user privacy and prevent unauthorised access to sensitive information.

**10.2. Case Study 2: Smart Public Transportation System.** Scenario: A sophisticated public transport system employs cloud-based solutions to optimise bus and railway routes, schedules, and fleet management.

#### Challenges.

- **Data Residency:** Concerning the physical location of data stored in the cloud and assuring conformance with data residency regulations, the transport authority confronts challenges.
- **Proprietorship Concerns:** The authority must clarify data possession and control, ensuring that an agency's transportation retains power over the data.

### Implementation Strategies.

- **Data Governance:** The transportation authority implements a robust data governance framework to define data ownership, storage, and access policies, ensuring data residency requirements are met.
- **Transparent Agreements:** The transport authority establishes clear agreements with the cloud service provider regarding data custody and control, granting it full administrative rights over its data.

**Results.** By addressing concerns regarding data residency and ownership, the intelligent public transport system obtains the confidence of stakeholders and ensures compliance with applicable regulations. The effective cloud-based solution optimises conveyance services, enhancing passenger satisfaction and decreasing operational expenses. Case studies and real-world examples illustrate the practical implications of cloud-based security measures in intelligent transportation applications. Cloud-based solutions are indispensable for optimising transportation systems and protecting sensitive data by addressing data privacy, security, and ownership challenges. These examples provide transportation authorities, industry stakeholders, and policymakers with valuable guidance towards adopting cloud-based solutions responsibly and securely while safeguarding data and privacy in intelligent transportation.

**11. Discussion.** The discussion section emphasises the most important considerations and strategies for data and privacy protection in cloud-based intelligent transportation applications. It examines the role of cloud computing in addressing data security and privacy issues, but it is crucial to evaluate the efficacy and limitations of these solutions.

While encryption and access control mechanisms provided by cloud platforms can improve data security, it is essential to observe that encryption alone does not provide absolute security. Encryption algorithms have occasionally been compromised or inadequately implemented, resulting in data intrusions. In addition, access control mechanisms are only effective if they are appropriately configured and managed, as missed configurations or vulnerabilities can result in unauthorised access.

There are disadvantages to cloud storage and backup solutions for data protection. Organisations must choose cloud service providers with a strong security focus and robust data protection measures with care. Concerns persist regarding data residency and proprietorship, as organisations may have limited control over the physical location of their data and the permissions to access and administer it.

Even though cloud platforms provide threat detection and intrusion prevention measures, it is essential to recognise that these systems are not infallible. Complex cyber attacks can still circumvent detection mechanisms, and new vulnerabilities may emerge as technologies evolve. Continuous monitoring, regular updates, and proactive security measures are essential to remain abreast of evolving threats. The protection of privacy in cloud-based intelligent transportation applications is a challenging problem. Although anonymisation and pseudonymisation techniques can aid in protecting user identities, it is essential to recognise that re-identification attacks and privacy breaches can still occur. Privacy by design principles are crucial, but their successful implementation requires a comprehensive understanding of privacy risks and impact assessments.

Cloud platforms' data governance and compliance frameworks can aid organisations in meeting regulatory requirements. However, organisations must also be accountable for comprehending and abiding by applicable data protection laws and regulations. Compliance should not rely solely on cloud service providers because organisations are ultimately responsible for the security and confidentiality of their data. In intelligent transportation applications, it is essential to recognise cloud-based solutions' potential obstacles and limitations for data security and privacy. Organisations must thoroughly evaluate the risks and benefits of these obstacles. Adopting a comprehensive approach to security that incorporates many levels of protection, frequent audits, and continuing training is suggested to limit risks and preserve the integrity and privacy of data.

Even though cloud-based solutions provide beneficial security and privacy features, they are not fail-safe. In intelligent transportation applications, organisations must exercise caution and implement additional safeguards to address the limitations and potential risks associated with cloud computing. Continuous evaluation, proactive security measures, and comprehensive security strategies are essential for sustaining data security and privacy in a constantly changing environment. Several limitations were encountered during the research for "Protecting Data and Privacy: Cloud-based Solutions for Intelligent Transportation," which affected the study's scope and findings. Data availability complicated access to exhaustive and up-to-date information on real-world cloud-based intelligent transportation implementations, potentially resulting in data gaps. Despite the systematic

approach, the extensive literature on data security, privacy, and cloud-based solutions in smart transportation made it difficult to identify all relevant studies, resulting in unintentional omissions. In addition, the analysis was incomplete due to cloud service providers' limited disclosure of security measures.

Due to the rapid evolution of cloud computing and intelligent transportation systems, it was possible that new advancements would emerge during research that could not be completely incorporated. When discussing real-world case studies or industry-specific implementations, the amount of specific information that could be disclosed was also influenced by ethical considerations surrounding data privacy and confidentiality.

Investigating the implementation challenges organisations face when deploying cloud-based solutions, conducting a comparative analysis of the security features of different cloud service providers, and understanding user perceptions and acceptance of data security and privacy in intelligent transportation applications are examples of research gaps and areas for further investigation. In addition to examining the challenges of complying with data protection regulations across jurisdictions and international privacy frameworks, it is vital to conduct long-term assessments of cloud-based security measures and their effectiveness over extended periods.

Addressing these research voids will contribute to a deeper comprehension of data security and privacy in cloud-based intelligent transportation, allowing for more informed decision-making and developing robust security strategies in this swiftly evolving field. It will also pave the way for enhancing user confidence, data security, and the development of intelligent transportation applications over time.

**12. Conclusion.** As intelligent transportation systems continue to evolve, it becomes increasingly important to protect the security and privacy of data. Cloud-based solutions provide a robust framework for effectively addressing these concerns. Organizations can protect sensitive data and safeguard user privacy by implementing encryption, access controls, threat detection mechanisms, and privacy protection measures. Adopting these cloud-based solutions will facilitate the extensive adoption of intelligent transportation applications while infusing users and stakeholders with confidence.

#### REFERENCES

- [1] Abd-El-Atty, B., Ilyyasu, A. M., Alaskar, H., Abd El-Latif, A. A. (2020). A robust quasi-quantum walks-based steganography protocol for secure transmission of images on cloud-based E-healthcare platforms. *Sensors*, 20(11), 3108.
- [2] Alam, T. (2021). Cloud-based IoT applications and their roles in smart cities. *Smart Cities*, 4(3), 1196-1219
- [3] Alkharji, L., De, S., Rana, O., and Perera, C. (2023). Semantics-based privacy by design for Internet of Things applications. *Future Generation Computer Systems*, 138, 280-295.
- [4] Butpheng, C., Yeh, K. H., Xiong, H. (2020). Security and privacy in IoT-cloud-based e-health systems—A comprehensive review. *Symmetry*, 12(7), 1191.
- [5] Cavoukian, A. (2010). Privacy by design: the definitive workshop. *Identity in the Information Society*, 3(2), 247-251.
- [6] Cheng, L., Liu, F., Yao, D. (2017). Enterprise data breach: causes, challenges, prevention, and future directions. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 7(5), e1211.
- [7] Conte, R., Sansone, F., Tonacci, A., Pala, A. P. (2022). Privacy-by-Design and Minimization within a Small Electronic Health Record: The Health360 Case Study. *Applied Sciences*, 12(17), 8441
- [8] Das, D., Banerjee, S., Chatterjee, P., Ghosh, U., Biswas, U. (2023). Blockchain for Intelligent Transportation Systems: Applications, Challenges, and Opportunities. *IEEE Internet of Things Journal*.
- [9] Gholamhosseinian, A., and Seitz, J. (2021). Vehicle classification in intelligent transport systems: An overview, methods and software perspective. *IEEE Open Journal of Intelligent Transportation Systems*, 2, 173-194
- [10] Gohar, A., and Nencioni, G. (2021). The role of 5G technologies in a smart city: The case for intelligent transportation system. *Sustainability*, 13(9), 5188.
- [11] Guerrero-Ibanez, J. A., Zeadally, S., Contreras-Castillo, J. (2015). Integration challenges of intelligent transportation systems with connected vehicle, cloud computing, and internet of things technologies. *IEEE Wireless Communications*, 22(6), 122-128.
- [12] Gunduz, M. Z., Das, R. (2020). Cyber-security on smart grid: Threats and potential solutions. *Computer Networks*, 169, 107094.
- [13] Haque, A. B., Bhushan, B., Dhiman, G. (2022). Conceptualising smart city applications: Requirements, architecture, security issues, and emerging trends. *Expert Systems*, 39(5), e12753.
- [14] Hon, W. K., Millard, C., Walden, I. (2011). The problem of 'personal data' in cloud computing: what information is regulated?—the cloud of unknowing. *International Data Privacy Law*, 1(4), 211-228.
- [15] Hoofnagle, C. J., Van Der Sloot, B., Borgesius, F. Z. (2019). The European Union general data protection regulation: what it is and what it means. *Information Communications Technology Law*, 28(1), 65-98.
- [16] Hussain, M. M., Alam, M. S., Beg, M. S. (2019). Fog computing model for evolving smart transportation applications. *Fog and Edge Computing: Principles and Paradigms*, 22(4), 347-372.

- [17] Joshi, J. B., Bertino, E., Latif, U., Ghafoor, A. (2005). A generalized temporal role-based access control model. *IEEE transactions on knowledge and data engineering*, 17(1), 4-23
- [18] Kaluarachchi, Y. (2022). Implementing data-driven smart city applications for future cities. *Smart Cities*, 5(2), 455-474.
- [19] Latha, V. P., Reddy, N. S., Babu, A. S. (2023). Optimizing Scalability and Availability of Cloud Based Software Services Using Modified Scale Rate Limiting Algorithm. *Theoretical Computer Science*, 943, 230-240.
- [20] Lee, I., Lee, K. (2015). The Internet of Things (IoT): Applications, investments, and challenges for enterprises. *Business Horizons*, 58(4), 431-440.
- [21] Masud, M., Gaba, G. S., Choudhary, K., Alrobaea, R., Hossain, M. S. (2021). A robust and lightweight secure access scheme for cloud based E-healthcare services. *Peer-to-peer Networking and Applications*, 14(5), 3043-3057.
- [22] Mehta, P., Gupta, R., Tanwar, S. (2020). Blockchain envisioned UAV networks: Challenges, solutions, and comparisons. *Computer Communications*, 151, 518-538.
- [23] Mollah, M. B., Zhao, J., Niyato, D., Guan, Y. L., Yuen, C., Sun, S. et al (2020). Blockchain for the internet of vehicles towards intelligent transportation systems: A survey. *IEEE Internet of Things Journal*, 8(6), 4157-4185.
- [24] Mueller, S., Taylor, C. R., Mueller, B. (2022). *Managing change related to consumer privacy laws: targeting and personal data use in a more regulated environment*. In Media and Change Management Creating a Path for New Content Formats, Business Models, Consumer Roles, and Business Responsibility. Cham Springer International Publishing
- [25] Narwal, P., Duhan, N., Bhatia, K. K. (2022). Image Systems and Visualisations 3. *Multimedia Computing Systems and Virtual Reality*, 45-78.
- [26] Oprescu, A. M., Mir'ó-Amarante, G., García-Díaz, L., Rey, V. E., Chimenea-Toscano, A., MartínezMartínez, R., Romero-Ternero, M. C. (2022). Towards a data collection methodology for Responsible Artificial Intelligence in health: A prospective and qualitative study in pregnancy. *Information Fusion*, 83, 53-78.
- [27] Pandeewari, N., Kumar, G. (2016). Anomaly detection system in cloud environment using fuzzy clustering based ANN. *Mobile Networks and Applications*, 21, 494-505.
- [28] Patel, A., Taghavi, M., Bakhtiyari, K., Júnior, J. C. (2013). An intrusion detection and prevention system in cloud computing: A systematic review. *Journal of Network and Computer Applications*, 36(1), 25-41.
- [29] Pérez, J., León, J., Castilla, Y., Shahrabadi, S., Anjos, V., Ad'ao, T., et al (2023). A cloud-based 3D real-time inspection platform for industry: a case-study focusing automotive cast iron parts. *Procedia Computer Science*, 219, 339-344
- [30] Ranchal, R., Bastide, P., Wang, X., Gkoulalas-Divanis, A., Mehra, M., Bakthavachalam, S., et al (2020). Disrupting healthcare silos: Addressing data volume, velocity and variety with a cloud-native healthcare data ingestion service. *IEEE Journal of Biomedical and Health Informatics*, 24(11), 3182-3188.
- [31] Rimal, B. P., Jukan, A., Katsaros, D., Goeleven, Y. (2011). Architectural requirements for cloud computing systems: an enterprise cloud approach. *Journal of Grid Computing*, 9, 3-26.
- [32] Roy, P. P. (2020). A high-level comparison between the nist cyber security framework and the iso 27001 information security standard. In *2020 National Conference on Emerging Trends on Sustainable Technology and Engineering Applications (NCETSTEA)* (pp. 1-3). IEEE.
- [33] Saranya, N., Sakthivadivel, M., Karthikeyan, G., Rajkumar, R. (2023). Securing the Cloud: An Empirical Study on Best Practices for Ensuring Data Privacy and Protection. *International Journal of Engineering and Management Research*, 13(2), 46-49
- [34] Samantha, F. H., Azam, S., Yeo, K. C., Shanmugam, B. (2020). A systematic literature review on privacy by design in the healthcare sector. *Electronics*, 9(3), 452.
- [35] Singh, B., Gupta, A. (2015). Recent trends in intelligent transportation systems: a review. *Journal of Transport Literature*, 9, 30-34.
- [36] Sood, S. K. (2021). Smart vehicular traffic management: An edge cloud centric IoT based framework. *Internet of Things*, 14, 100140.
- [37] Wang, Y., Su, Z., Zhang, N., Xing, R., Liu, D., Luan, T. H., Shen, X. (2022). A survey on metaverse: Fundamentals, security, and privacy. *IEEE Communications Surveys Tutorials*
- [38] Zhao, S., Miao, J., Zhao, J., Naghshbandi, N. (2023). A comprehensive and systematic review of the banking systems based on pay-as-you-go payment fashion and cloud computing in the pandemic era. *Information Systems and e-Business Management*, 1-29

*Edited by:* Zhenling Liu

*Special Issue on:* Cloud Computing for Intelligent Traffic Management and Control

*Received:* Jun 30, 2023

*Accepted:* Aug 21, 2023



## IMPROVING BERT MODEL ACCURACY FOR UNI-MODAL ASPECT-BASED SENTIMENT ANALYSIS TASK

**Abstract.** Techniques and methods for examining users' feelings, emotions, and views in text or other media are known as "sentiment analysis," this phrase is used frequently. In many areas, including marketing and online social media, analysis of user and consumer opinions has always been essential to decision-making processes. The development of new methodologies that concentrate on analysing the sentiment associated with specific product characteristics, such as aspect-based sentiment analysis (ABSA), was prompted by the need for a deeper understanding of these opinions. Despite the growing interest in this field, some misunderstanding exists about ABSA's core ideas. Even though sentiment, affect, emotion, and opinion refer to various ideas, they are frequently used synonymously. This ambiguity commonly causes user opinions to be analysed incorrectly. This work provides an overview of ABSA and the issue of overfitting. Following this analysis, we improved the model by enhancing the accuracy and F1 score of the existing model by fine-tuning the technique. Our model outperformed the others, achieving the best results for the restaurant dataset with an 85.02 accuracy and a 79.19 F1 score, respectively.

**Key words:** Aspect-based sentiment analysis, Sentiment analysis, Natural language processing, Online Social networks, Opinion Mining.

**1. Introduction.** ABSA is a technique used in Natural Language Processing (NLP) to ascertain a text's attitude towards a particular quality or attribute of a good, service, or institution. It entails determining the sentiment polarity (neutral, positive or negative) for a specific interest feature, such as a smartphone's battery life or a cuisine's flavour. This type of analysis is essential in several industries, including marketing, product development, and customer service, as it provides valuable insights into customer opinions and preferences. ABSA frequently includes NLP techniques, such as sentiment analysis (SA), named object identification, and topic modelling [27].

Meanwhile, unimodal SA analyses unimodal data, such as text from [25], to determine people's attitudes towards various topics or products. Online platforms like Twitter and Facebook have become well-liked methods to share and post views, including text, pictures, and audio, as social media and web technology have grown in popularity.

Analysis of the sentiment of a sentence towards particular opinion objectives is the job involved in ABSA. This includes identifying the polarity of sentiment (positive, neutral or negative) for each target mentioned in the statement, such as "SERVICE" and "FOOD," which would be negative and positive, respectively [24].

For example, consider a real-life scenario where you work for an e-commerce company that analyses customer reviews of a new smartphone. Your task is to perform ABSA on the reviews to obtain information about the aspects customers discuss and their sentiments towards each element. In the given study, "I like the design of the phone, but the battery life is a major disappointment," there are two aspects - A1: design and A2: battery life. Design is a positive aspect, while battery life is a negative aspect.

Previous studies have primarily relied on RNNs to model text, which suffers from complex parallelization, memory and computation requirements, and the inability to capture long-term dependencies due to truncated backpropagation through time (BPTT). Although LSTMs can maintain long-range information and partially solve the issue of disappearing gradients, they need a lot of training data. The label unreliability problem, particularly for neutral sentiment, a fuzzy condition that makes learning more difficult, is another difficulty that has been ignored in earlier studies.

Targeted sentiment classification involves identifying sentiment towards a specific target entity or aspect in text, but label unreliability is a common challenge in this task [25]. This refers to the inconsistency or inaccuracy of the labelled data used to train a sentiment classifier due to subjective or ambiguous sentiment labels assigned by different annotators. One approach uses inter-annotator agreement metrics to address this issue and filter

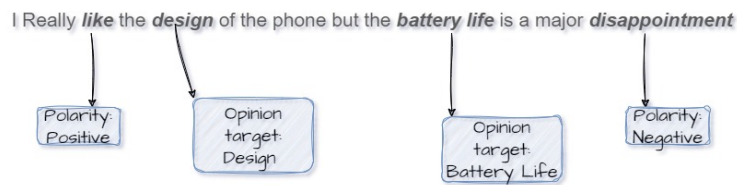


Fig. 1.1: Exemplary ABSA Statement

out instances where annotators disagree. In contrast, another uses active learning to select informative or uncertain illustrations for annotation and update the model iteratively. Semi-supervised learning can also leverage improving both labelled and unlabeled data classifier performance and robustness.

To enhance the effectiveness of the essential BERT model for sentiment classification, especially for fuzzy labels, this research suggests a model that uses label smoothing regularisation [24]. Our model outperforms basic BERT when used with pre-trained BERT on three benchmark datasets, demonstrating competitive performance and lightweight properties. The paper’s key contributions include addressing overfitting issues, applying pre-trained BERT, and using fine-tuning modified BERT SPC model to overcome overfitting in target and context words.

**1.1. Motivation.** This study’s primary objective is to assess sentiment in opinionated text where users provide feedback about products, services, or events through writing reviews. By defining terms like emotion and affective opinions, an effective ABSA model must be able to recognise different sentimental dimensions. The literature review section highlights the need for a commonly agreed-upon definition of fundamental ABSA concepts and appropriate metrics for each measurement. Positivity about a product differs from expressing sentiment towards it and thus demands various measurement and evaluation methods. While computer scientists may conflate feelings and opinions, this is not true in other fields such as social sciences, psychology, marketing and communications, and neuroscience. This divergence poses a genuine obstacle to creating reliable and practical real-world applications. Defining the properties of the document or media being analysed is a crucial issue that ABSA models must handle. Review analysis primarily concentrates on user-generated text. However, contemporary websites and social media platforms provide several tools and features that let consumers write in-depth and elaborate reviews of goods or services. In addition to the main text, users can provide ratings on a five-star scale, add tags to categorize their thoughts, use emoticons to convey their sentiments, and more. Traditional ABSA techniques typically consider only some of these elements.

**2. Related Work.** ABSA is a powerful NLP technique that indicates a sentence’s polarity for a specific aspect term. Giuseppe D’Aniello et al. (2022) [8] have provided an overview of the latest ABSA methodologies and approaches, highlighting the significant challenges of recent developments in this area. Their research introduces a new reference model called KnowMIS-ABSA, which can be extended to include reviews and other ABSA features in future work. Bin Liang et al. (2022) [9] introduced Sentic GCN, a graph convolutional network that utilizes emotional relationships of text for a specific aspect by using SenticNet. They use dynamic knowledge to enhance sentence dependency graphs and build unique chart neural networks. Haiyan Wu et al. (2022) [20] have introduced an attention network with phrase dependency for the ABSA task (PD-RGAT). The phrase dependency graph is created using this relational graph attention network by combining directed dependency edges and phrase information. GloVe and BERT, two distinct pre-training models, were tested and obtained results comparable to other baseline models. Kai He et al. (2022) [14] the meta-based MSM self-training system, which has a meta-weighter, was proposed by the authors. According to the authors, a neural system with effective learning control and appropriate symbolic representation choice can produce a generalizable model. They use MSM to train a teacher model to create in-domain knowledge; a student model subsequently uses it for supervised learning. Zigu Zhao et al. (2022) [26] a graph convolutional network with several weighting strategies was developed for aspect-based sentiment analysis. To fully utilize BERT, they introduced a dynamic weight alignment technique. They also devised an aspect-aware weight mechanism

to govern message propagation to the aspect. They introduced an aspect-oriented loading layer to limit the negative impacts of words unrelated to the aspect term. Finally, they fused high-order semantic and grammatical information via multi-head self-attention to forecast premium aspect-specific representations. Li Yang et al. (2022) [22] developed the Cross-Modal Multitask Transformer (CMMT), a multitask learning framework that includes two additional activities to practise aspect- and emotion-aware intra-modal examples. They also present a Text-Guided Cross-Modal Communication Module to dynamically regulate the contributions of visual information to each word's representation in the inter-modal interaction. Shi Feng et al. (2022) [11] proposed a new ABSA model called AG-VSR that relies on aspect representations rather than sentence representations and updates them with GCNs. To overcome the dependency on the integrity of the dependency tree and retain global sentence information, they used Attention-assisted Graph-based Representation (A GR) and Variational Sentence Representation (VSR). The GCN module generates a GR by modifying a dependency tree with the attention mechanism. Lastly, [4] Anan Dai et al. (2022) proposed a human cognition-based approach to It entails instruction in everything from sentence grammar to word meaning. To capture words' structural and broad semantics, they created a dual-channel semantic learning graph convolutional network (GCN). They then carried out a syntactic GCN to learn the syntactic structure of sentences. Their approach aligns with human processing practices and provides a meaningful interpretation of the given sentence.

**3. Sentiment Analysis.** Opinion Mining (OM) and Sentiment Analysis have emerged as established research fields over the last twenty years, finding widespread use in various applications, including commercial ones. The definitions of OM and SA and any distinctions between opinion, sentiment, emotion, affect, and related concepts continue to be unclear despite the enormous advancement and innovation in these disciplines [7]. Many academics contend that these notions are separate and call for various methodologies that result in multiple conclusions, even though some see this as a dispute over language with no genuine difference between these concepts. Despite the fact that both OM and SA seek to grasp the subjectivity expressed in a text, the exact definitions of opinion and sentiment are up for debate among scholars.

Sentiment analysis involves analyzing opinions at various levels, including the sentence, document, aspect, and concept levels. The Document-level analysis aims to learn the polarity of an entire document, such as a review or news article, using a single positive or negative score to condense the contradiction of several words. Approaches based on lexicons are frequently employed in document-level analysis [2], determining the polarity of each term before combining them to categorise the overall sentiment of the material. However, some researchers argue that considering the importance of each word could improve classification accuracy [3].

Sentence-level SA aims to understand a single sentence's negative or positive opinion. This method is frequently used to evaluate the sentiment of an entire document by combining the polarity scores of each sentence. While sentiment analysis at the text or sentence level offers helpful insights into users' perceptions of particular entities, it does not reveal the specific elements, features, or aspects that influence user opinions. Aspect-level sentiment analysis is required for this type of analysis. This approach involves identifying each aspect's opinionated sentences, entity categories, elements, linguistic expressions, and polarities. Concept-level SA, on the other hand, uses semantic analysis methods to find and examine concepts in the text, allowing computers to comprehend natural language and emotions at a deep level. With encouraging outcomes, SenticNet is a lexical resource that assigns mood and emotion tags to facilitate concept-level SA.

**3.1. Aspect-Based Sentiment Analysis (ABSA).** The goal of ABSA, a technique for natural language processing, is to extract sentiment information from text based on various characteristics or attributes of a given subject or entity. Sentiment identification, the next stage in ABSA, requires calculating a sentiment score (regarding orientation or polarity) for every aspect obtained from the text. The score typically falls within a mathematical range, such as a five-point scale or a decimal number between [-1, +1], where -1 was highly damaging, and +1 was highly positive. It expresses the degree of positivity or negativity regarding the element. The task is also known as sentiment classification if the only objective is to ascertain whether the sentiment is favourable or adverse [19]. A thorough analysis of contemporary aspect extraction methods is conducted, and they are divided into three primary categories: unsupervised, semi-supervised, and supervised procedures [23].

The unsupervised category includes techniques such as frequency or statistics-based heuristics, syntactic dependency path-based approaches, and rule-based techniques. The semi-supervised category comprises lexicon-based, dependency tree-based, and graph-based, for example, methods. Supervised techniques, on the other

hand, employ machine learning approaches such as random fields, SVM, decision trees, autoencoders and neural networks [13]. Many methods for identifying sentiment can be categorised into three groups: lexicon-based, machine learning-based, and hybrid methods [21]. Let me explain ABSA with a case study from real life.

Consider a client who wants to reserve a hotel for her future vacation. She begins her study online and pursues numerous evaluations of hotels in a specific tourist area. She discovers a hotel review: "Great amenities and location; however, the service was lacking. The staff was impolite and uncooperative. Although the breakfast was not up to standard, the hotel was clean and big. The location, facilities, service, staff, room, and breakfast are a few of the numerous characteristics listed in the review that may be analysed using ABSA in this case. Let's see how ABSA might apply here:

Location: The reviewer described it as "great"; the tone is favourable.

Amenities: The critic did not identify any problems. Thus, the sentiment is favourable.

Service: The reviewer expressed dissatisfaction with the service, calling it "rude," "disappointing," and "unhelpful."

Employees: The reviewer's comments about the employees being "rude" and "unhelpful" reflect the overall poor opinion.

I had a positive feeling because it was described as "clean" and "spacious" in the room evaluation.

Breakfast: The reviewer expressed dissatisfaction, calling it "subpar".

To help her comprehend that while the location and amenities are excellent, the service and breakfast components are unfavourable, ABSA assists in extracting sentiment information from various aspects of the evaluation. Based on the most critical factors to the client, this information can help her make an informed decision about whether or not to reserve the hotel.

**4. The Proposed Method.** The study's authors utilized the SemEval 2014 Task 4 Public dataset for data collection in the first step. Figure 2 depicts the flow of the proposed research, where the second step involves data pre-processing. The following step is feature selection, followed by training until the early stop prompt is received. The K-fold cross-validation (K=6) is then implemented in stage 2. If the accuracy is unsatisfactory, hyperparameter tuning is applied to the model for better results. The pre-trained models of the embedding layer played a vital role in achieving better results [18].

The embedding layer is a crucial component of neural network models in NLP. It converts textual data, such as words or sentences, into numerical vectors that the model can process [5]. The embedding layer maps each word or token in the input text to a high-dimensional embedding vector. The size of the embedding vector is typically smaller than the vocabulary size, making processing more efficient. Pre-trained embeddings trained on large datasets using techniques such as Word2Vec or GloVe captures the semantic and syntactic links between words it can be helpful in NLP tasks. During training, the embedding layer adjusts the embedding vectors based on the error calculated by subsequent neural network layers [16]. This enables the model to learn how to produce embedding vectors that faithfully reflect the task's incoming data. Using an embedding layer has dramatically improved the accuracy of text classification, sentiment analysis [1], machine translation, and other NLP tasks by allowing the model to process textual data and capture complex word relationships efficiently.

The BERT (Bidirectional Encoder Representations from Transformers) model is a pre-trained language model that generates contextualized word embeddings for various NLP tasks. The mathematical procedure for computing embeddings uses input tokens, a pre-trained language model with L transformer layers, and a hidden size H in the BERT embedding [12]. The BERT model function takes an input sequence of tokens S and outputs a series of concealed states.

$$H = \{h_1, h_2, \dots, h_n\}, \quad (4.1)$$

Our research utilized Bert Embedding throughout the experimentation phase. We extracted the final hidden state corresponding to that token to acquire the BERT embedding for a particular permit or sub-word, as outlined in equations 4.1 and 4.2. This remote state results after the input sequence passes through the L-th transformer layer. As a result, the BERT embedding of the  $i$ -th token can be expressed as equation 4.2.

$$\text{BERT}(S)_i = h_i \wedge (L) \quad (4.2)$$



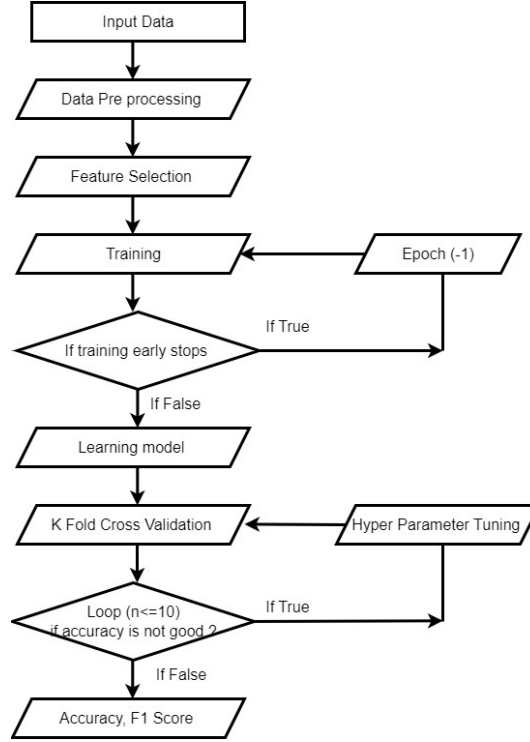


Fig. 4.1: The proposed Work

An alternative method of computing the BERT embedding involves computing the average or maximum hidden states of a sequence of tokens. Here, the set of indices  $I$  corresponds to the token sequence in  $S$ , and the BERT embedding of the sequence can be defined as equation 4.3.

$$\text{BERT}(S)_I = \text{mean} / \max (\{h_i \wedge (L) \mid i \text{ in } I\}) \quad (4.3)$$

To compute the BERT embedding, one can either take the final hidden state of a particular token or sub-word or average/take the maximum of the hidden forms of a sequence of tokens, as generated by the pre-trained BERT language model. The averaging or maximum operation is performed over the set of hidden states corresponding to the tickets in  $I$ .

**5. Experimental Setup.** Our experimentation used the Intel (R) Xeon (R) CPU E3-1225v5@3.31GHz Processor, 32 GB RAM, and a 64-bit-based processor.

**5.1. Dataset.** The SemEval-2014 Task 4 Subtask 2 dataset is intended for academics and programmers working on sentiment analysis and natural language processing tasks. The collection benefits academics who are building and testing algorithms for classifying the emotions of brief messages, like tweets. Researchers should utilise this data set because it has been annotated by qualified individuals with sentiment polarity tags. This indicates that the dataset offers robust ground truth data necessary for creating and testing precise sentiment analysis algorithms. The main reason for choosing the SemEval 2014 Task 4 subtask 2 dataset is because the set of data is a popular baseline data for evaluating sentiment analysis algorithms so that authors can compare the work to a substantial body of existing research and a standard assessment process. In general, using the SemEval-2014 Task 4 Subtask 2 dataset can assist in creating and assessing cutting-edge sentiment analysis algorithms, comparing the results to an accepted standard, and contributing to the larger natural language processing research tasks.

Table 5.1: Aspect words classified into positive, negative, and neutral polarities are present in all three datasets’ train/test sets.

Dataset		Instances	Positive	Negative	Neutral
LAPTOP 14	Train	2,328	994	870	464
	Test	638	341	128	169
RESTAURANT 14	Test	1,106	728	182	196
	Train	3,608	2,164	807	637
Twitter 14	Train	6,248	1561	1,560	3,127
	Test	692	173	173	346

To Perform our experimentation work, we employed the SemEval 2014 Task 4 Subtask 2 dataset [18]. The dataset comprises three domain datasets: Laptop, Twitter, and Restaurant. Each dataset contains separate train and test files.

This study includes experimentation using both a training testing set and a training training set to demonstrate that overfitting is a concern with the SemEval 2014 Task 4 dataset. When a higher epoch value is set, the early stop occurs during the training stage.

The number of times a model is trained on the full dataset is called an "epoch" during the training process [15]. Increasing the epoch value may allow the model to learn more complex patterns within the data. Still, it can also result in overfitting, where the model becomes too specialized in the training data, leading to poor performance on new data.

Early stopping methods are used during training to monitor a validation measure like accuracy or loss to improve the model’s performance and prevent overfitting. The activity is terminated if the validation measure subsequently ceases to improve. In some circumstances, setting the highest epoch value may be suitable. Still, it’s crucial to consider dataset size and model complexity when choosing the correct number of epochs. Early halting strategies that depend on validation measures work better in general.

To combat overfitting, we used the K-fold cross-validation approach. This method involves dividing the data into K similarly sized "folds," training the model on K-1 folds, and then testing it on the residual fold [10]. The process is repeated K times to get an overall assessment of the model’s performance, with each fold serving as the test set once.

Cross-validation can provide a more accurate estimate of the model’s performance than utilising a single train-test split to prevent overfitting by ensuring that the model is tested on a variety of data [28]. There are various types of k-fold cross-validation, such as stratified k-fold, which preserves the class distribution in each fold. Leave-one-out cross-validation is another variant, where each instance is used as the test set once, and K is set to the number of instances in the dataset.

Overall, K-fold cross-validation helps evaluate a model’s performance and can identify potential issues such as overfitting. In mathematical notation [6, 17], the formula for K-fold cross-validation can be expressed as follows in equation 5.1: Let  $X = \{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}$  be a dataset of size  $N$ , and  $K$  be the number of folds. For each fold  $i$  in  $K$ , let

$$x_i = \{(x_{i_1}, y_{i_1}), (x_{i_2}, y_{i_2}), \dots, (x_{i_{N_i}}, y_{i_{N_i}})\} \quad (5.1)$$

be the  $i$ -th fold of size  $N_i$ , such that  $N = k \cdot N_i$ . Let

$$x_{\text{train}} = \{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\} - X_i \quad (5.2)$$

The ( $X$  train) dataset was used to train the model, and the parameter vector theta was obtained to evaluate  $X$  test and get the predicted values ( $y$  pred). The evaluation metric(s) used, such as mean squared error or accuracy, were calculated between ( $y$  pred) and the actual values ( $y$  test) in ( $X$  test). This process was repeated for all  $i$  in  $K$ , using equation 5.2 to calculate the evaluation metric(s) each time.

The K-fold cross-validation technique was used to test the BERT-SPC model on three datasets, and the results are presented in Table 5.2. The restaurant dataset outperformed the other two datasets. During training,

Table 5.2: Without the Cross-Validation Technique

Dataset	Training	Testing	Training	Training
	Accuracy	F1	Accuracy	F1
Laptop	0.8058	0.7569	0.8990	0.8885
Twitter	0.7356	0.7011	0.9190	0.9180
Restaurant	0.8502	0.7919	0.9475	0.9224

Table 5.3: With K Fold Cross Validation Technique (K=6)

Dataset	Training	Testing	Training	Training
	Accuracy	<b>F1</b>	<b>Accuracy</b>	<b>F1</b>
Laptop	0.8058	0.7569	0.8990	0.8885
Twitter	0.7356	0.7011	0.9190	0.9180
Restaurant	0.8502	0.7919	0.9475	0.9224

the authors observed that the "Early Stop" prompt occurred in every epoch. To address this, they implemented the K Cross-Validation technique with K=6. The overall performance estimate was obtained by averaging the K evaluation metric(s) over all folds.

Table 5.3 displays the results obtained using the K Cross-Validation Technique with k=6, which led to improved outcomes. This technique addressed the issue of overfitting, and the K Cross-Validation technique performed better than other approaches.

**5.2. Model Comparison.** A comparison of the proposed model (BERTSPC with K fold cross-validation) and (baseline research work) is shown in Table 5.4. The authors improved the performance of the BERTSPC model with the hyperparameters tuning technique. The authors encountered the overfitting issue during experimental work. To address this issue, they implemented the K Fold cross-validation technique, which yielded the best results for the BERTSPC model. The authors also experimented with several cross-validation techniques, such as the Stratified Cross-validation technique, Leave One Out Cross-validation, and Group K Fold Validation, on the SemEval 2014 Dataset and achieved comparable results using the Leave One Out and Group K Fold techniques. However, the K Fold Cross-validation technique produced the best results in the authors' experimentation.

Figure 3,4 and 5 shows the Accuracy and F1 Score of various baseline models for the Twitter, Restaurant and Laptop Dataset, respectively. From Figure 3 for the Twitter dataset in the previous BERTSPC model, the accuracy is 73.55, and in our proposed model, we achieved 0.01 high precision compared to the previous model. From Figure 3, we can see ATAE-LSTM and IAN are not plotted because, in previous studies, these models were not tested on this dataset. From Figure 4 for the [27]Restaurant dataset in the previous BERTSPC model, the accuracy is 84.46, and in our proposed model, we achieved 0.56 high precision compared to the previous model. The laptop dataset's highest level of accuracy is shown in Figure 5. From Figure 5, the accuracy of the laptop dataset in the last BERTSPC model is 73.55, and in our proposed model, we achieved 1.59 high precision compared to the previous model.

**6. Conclusion.** This study focuses on absa, which utilizes Attention-based Encoders for Target and Context. We used a pre-trained BERT model for this task to achieve the best possible outcomes. Through the process of hyperparameter tuning, we were able to enhance the accuracy of the Bert Spc model. However, during experimentation, we identified an overfitting issue addressed by implementing the Cross-validation technique. Our findings indicate that the K Cross-validation approach with K=6 produced the best results. Moving forward, we plan to apply this approach to real-world datasets. While this is a unimodal problem, we intend to extend our policy to multimodal issues.

Table 5.4: Baseline model findings are taken from articles that have already been published, (-) means not available

Models	Twitter		Restaurant		Laptop	
	Acc.	F1	Acc.	F1	Acc.	F1
TD LSTM	0.7080	0.6900	0.7563	-	0.6813	-
ATAE LSTM	-	-	0.7720	-	0.6870	-
IAN	-	-	0.7860	-	0.7210	-
RAM	0.6936	0.6730	0.8023	0.7080	0.7449	0.7135
FEATURE-BASED SVM	0.6340	0.6330	0.8016	-	0.7049	-
REC-NN	0.6630	0.6590	-	-	-	-
MEMNET	0.6850	0.6691	0.7816	0.6583	0.7033	0.6409
AEN- GloVe w/o PCT	0.7066	0.6907	0.8017	0.7050	0.7272	0.6750
AEN-GloVe w/o MHA	0.7124	0.6953	0.7919	0.7028	0.7178	0.6650
AEN-GloVe w/o LSR	0.7080	0.6920	0.8000	0.7108	0.7288	0.6869
AEN- GloVe BiLSTM	0.7210	0.7042	0.7973	0.7037	0.7312	0.6980
AEN GloVe	0.7283	0.6981	0.8098	0.7214	0.7351	0.6904
BERT SPC	0.7355	0.7214	0.8446	0.7698	0.7899	0.7503
AEN BERT	0.7471	0.7313	0.8312	0.7376	0.7993	0.7631
<b>BERT SPC with K Fold Cross Validation (Proposed)</b>	<b>0.7356</b>	<b>0.7011</b>	<b>0.8502</b>	<b>0.7919</b>	<b>0.8058</b>	<b>0.7569</b>

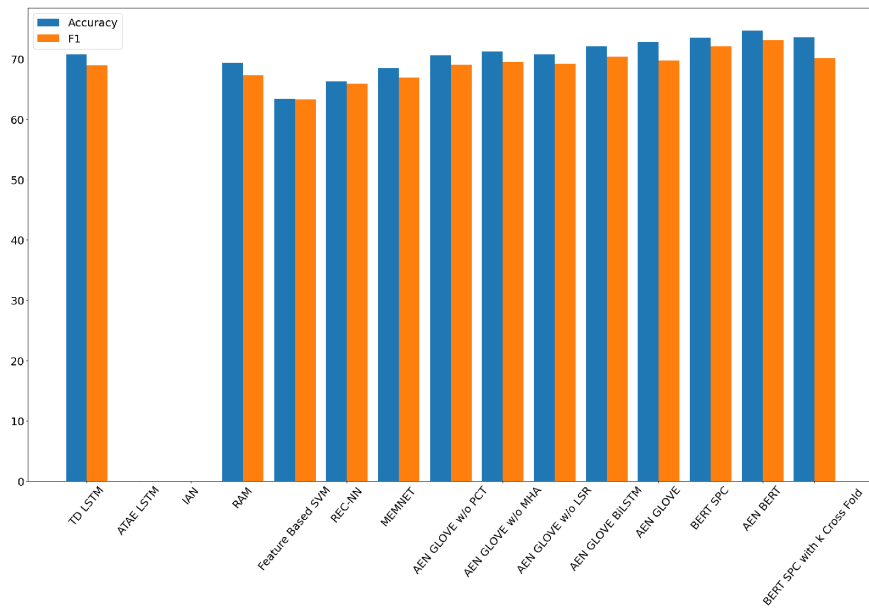


Fig. 5.1: Accuracy and F1 score on Twitter Data set

## REFERENCES

- [1] W. T. ALSHAMMARI AND S. ALHUMOUD, *Taqs: An arabic question similarity system using transfer learning of bert with bilstm*, IEEE Access, 10 (2022), pp. 91509–91523.
- [2] A. CHAUHAN AND R. MOHANA, *Implementing lda topic modelling technique to study user reviews in tourism*, in 2022 Seventh International Conference on Parallel, Distributed and Grid Computing (PDGC), IEEE, 2022, pp. 357–360.

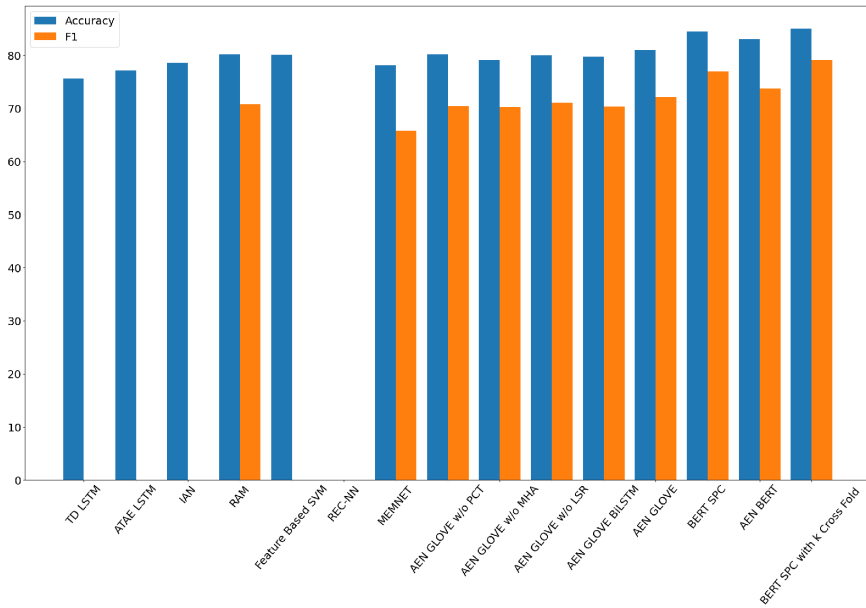


Fig. 5.2: Accuracy and F1 score on Restaurant Dataset

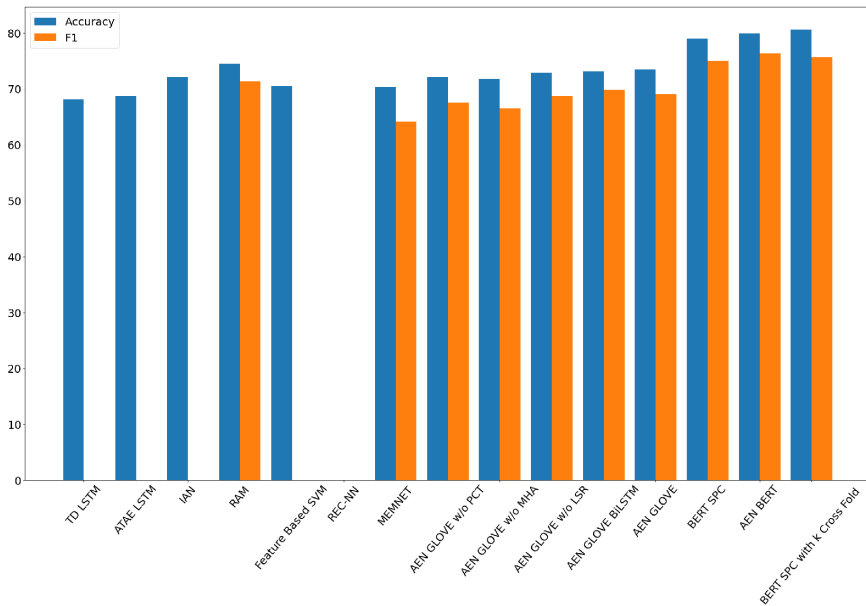


Fig. 5.3: Accuracy and F1 score on Laptop Dataset

- [3] A. CHAUHAN AND M. SOOD, *Sybil account detection in online social networks using statistical feature selection techniques*, in Futuristic Trends in Networks and Computing Technologies: Second International Conference, FTNCT 2019, Chandigarh, India, November 22–23, 2019, Revised Selected Papers 2, Springer, 2020, pp. 605–618.
- [4] A. DAI, X. HU, J. NIE, AND J. CHEN, *Learning from word semantics to sentence syntax by graph convolutional networks for aspect-based sentiment analysis*, International Journal of Data Science and Analytics, 14 (2022), pp. 17–26.
- [5] Z. DING, H. LI, W. SHANG, AND T.-H. P. CHEN, *Can pre-trained code embeddings improve model performance? revisiting the use of code embeddings in software engineering tasks*, Empirical Software Engineering, 27 (2022), p. 63.

- [6] J. D. DOMINGO, R. M. APARICIO, AND L. M. G. RODRIGO, *Cross validation voting for improving cnn classification in grocery products*, IEEE Access, 10 (2022), pp. 20913–20925.
- [7] G. DUBEY, H. P. SINGH, K. SHEORAN, G. DHAND, AND P. MALIK, *Drug review sentimental analysis based on modular lexicon generation and a fusion of bidirectional threshold weighted mapping cnn-rnn*, Concurrency and Computation: Practice and Experience, (2023), p. e7512.
- [8] G. D’ANIELLO, M. GAETA, AND I. LA ROCCA, *Knowmis-absa: an overview and a reference model for applications of sentiment analysis and aspect-based sentiment analysis*, Artificial Intelligence Review, 55 (2022), pp. 5543–5574.
- [9] ———, *Knowmis-absa: an overview and a reference model for applications of sentiment analysis and aspect-based sentiment analysis*, Artificial Intelligence Review, 55 (2022), pp. 5543–5574.
- [10] N. ELARABY, S. BARAKAT, AND A. REZK, *A conditional gan-based approach for enhancing transfer learning performance in few-shot hcr tasks*, Scientific Reports, 12 (2022), p. 16271.
- [11] S. FENG, B. WANG, Z. YANG, AND J. OUYANG, *Aspect-based sentiment analysis with attention-assisted graph and variational sentence representation*, Knowledge-Based Systems, 258 (2022), p. 109975.
- [12] X. FENG AND Y. ZENG, *Neural collaborative embedding from reviews for recommendation*, IEEE Access, 7 (2019), pp. 103263–103274.
- [13] J. HE, A. WUMAIER, Z. KADEER, W. SUN, X. XIN, AND L. ZHENG, *A local and global context focus multilingual learning model for aspect-based sentiment analysis*, IEEE Access, 10 (2022), pp. 84135–84146.
- [14] K. HE, R. MAO, T. GONG, C. LI, AND E. CAMBRIA, *Meta-based self-training and re-weighting for aspect-based sentiment analysis*, IEEE Transactions on Affective Computing, (2022).
- [15] T.-H. LI, H. ZHANG, L. FAN, H. WANG, AND Q. LIU, *Research on deep neural network model construction and overfitting*, in International Conference on Neural Networks, Information, and Communication Engineering (NNICE), vol. 12258, SPIE, 2022, pp. 88–93.
- [16] X. LI, X. FU, G. XU, Y. YANG, J. WANG, L. JIN, Q. LIU, AND T. XIANG, *Enhancing bert representation with context-aware embedding for aspect-based sentiment analysis*, IEEE Access, 8 (2020), pp. 46868–46876.
- [17] A. MAITY, S. GHOSH, S. KARFA, M. MUKHOPADHYAY, S. PAL, AND P. K. D. PRAMANIK, *Sentiment analysis from travellers’ reviews using enhanced conjunction rule based approach for feature-specific evaluation of hotels*, Journal of Statistics and Management Systems, 23 (2020), pp. 983–997.
- [18] M. PONTIKI, D. GALANIS, J. PAVLOPOULOS, H. PAPAGEORGIOU, I. ANDROUTSOPOULOS, AND S. MANANDHAR, *SemEval-2014 task 4: Aspect based sentiment analysis*, in Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014), Dublin, Ireland, Aug. 2014, Association for Computational Linguistics, pp. 27–35.
- [19] T. A. RANA AND Y.-N. CHEAH, *Aspect extraction in sentiment analysis: comparative analysis and survey*, Artificial Intelligence Review, 46 (2016), pp. 459–483.
- [20] H. WU, Z. ZHANG, S. SHI, Q. WU, AND H. SONG, *Phrase dependency relational graph attention network for aspect-based sentiment analysis*, Knowledge-Based Systems, 236 (2022), p. 107736.
- [21] K. XU, F. LI, D. XIE, AND D. JI, *Revisiting aspect-sentiment-opinion triplet extraction: Detailed analyses towards a simple and effective span-based model*, IEEE/ACM Transactions on Audio, Speech, and Language Processing, 30 (2022), pp. 2918–2927.
- [22] L. YANG, J.-C. NA, AND J. YU, *Cross-modal multitask transformer for end-to-end multimodal aspect-based sentiment analysis*, Information Processing & Management, 59 (2022), p. 103038.
- [23] Q. YANG, Z. KADEER, W. GU, W. SUN, AND A. WUMAIER, *Affective knowledge augmented interactive graph convolutional network for chinese-oriented aspect-based sentiment analysis*, IEEE Access, 10 (2022), pp. 130686–130698.
- [24] Q. YONG, C. CHEN, Z. WANG, R. XIAO, AND H. TANG, *Sgpt: Semantic graphs based pre-training for aspect-based sentiment analysis*, World Wide Web, (2023), pp. 1–14.
- [25] G. ZHAO, Y. LUO, Q. CHEN, AND X. QIAN, *Aspect-based sentiment analysis via multitask learning for online reviews*, Knowledge-Based Systems, (2023), p. 110326.
- [26] Z. ZHAO, M. TANG, W. TANG, C. WANG, AND X. CHEN, *Graph convolutional network with multiple weight mechanisms for aspect-based sentiment analysis*, Neurocomputing, 500 (2022), pp. 124–134.
- [27] Q. ZHONG, L. DING, J. LIU, B. DU, H. JIN, AND D. TAO, *Knowledge graph augmented network towards multiview representation learning for aspect-based sentiment analysis*, IEEE Transactions on Knowledge and Data Engineering, (2023).
- [28] Y. ZHONG, P. CHALISE, AND J. HE, *Nested cross-validation with ensemble feature selection and classification model for high-dimensional biological data*, Communications in statistics-simulation and computation, 52 (2023), pp. 110–125.

*Edited by:* Rajni Mohana

*Special Issue on:* Sentiment Analysis and Affective computing in Multimedia Data on Social Network

*Received:* Jul 25, 2023

*Accepted:* Sep 4, 2023



## DESIGN OF SENTIMENT ANALYSIS FRAMEWORK OF DIGITAL MEDIA SHORT TEXT BASED ON MULTI-PATTERN SENTIMENT LEXICON

**Abstract.** Along the continuous advancement of the network and the rise of digital media, the amount of data produced by the exponential explosion. And how to use these data to provide personalized services for users is one of the current research focuses. To address the issue of insufficient coverage in the current sentiment lexicon and the difficulty of constructing sentiment lexicon in specific fields, this study proposes a multi-modal emotional thesaurus. Semi-supervised learning is used to solve the problem of insufficient coverage of emotional thesaurus, and a semi-supervised classification algorithm is realized by using a large number of unlabeled sample data combined with a small number of labeled sample data. Optimized learning is used to solve the problem of difficult construction of emotional thesaurus in specific fields, the corresponding specific emotional thesaurus is constructed by adaptive adjustment of emotional word score, and finally the improved emotional thesaurus is used to build a digital media short text sentiment analysis framework. For testing, the NLPCC dataset was used in this study, Experiments show that the framework constructed in this study requires 87 iterations, a Recall value of 0.912, a F1 value of 0.753, and an average accuracy of 83.39%, all of which are better than the sentiment analysis framework without the use of multi-pattern sentiment lexicon. In the simulation experiment, the recognition accuracy reached 85.88%, which was 16.85%, 11.57% and 6.72% higher than the test scenarios using a single emotion thesaurus selected in this study. The above results show that the digital media short-text sentiment analysis framework built in this research based on multi-pattern sentiment lexicon can carry out short-text sentiment analysis more accurately and efficiently, so as to accurately analyze users' needs and provide customized services precisely.

**Key words:** Sentiment lexicon; Sentiment analysis; Semi-supervised learning; Optimization learning

**1. Introduction.** As the growth of the Internet, it has changed our life with more social media entering our daily life and more people beginning to evaluate the current hot events on digital media as a daily activity [1]. The increase of network users leads to the proliferation of data, which can provide reference information, but because the user's expression is not formal, the data will show confusion. Therefore, how to extract effective information from these data is one of the key types of research [4]. Text sentiment analysis (SA) technology in natural language processing technology can effectively process data to judge users' emotions, attitudes and expression of views [16]. The most basic semantic unit of SA is emotion words, and the collection of emotion words and their emotion scores and labels is sentiment lexicon (SL), through which the emotion polarity or emotion intensity of text can be analyzed and calculated [15]. Text sentiment analysis technology has achieved outstanding results in theoretical research and practical application, but with the development of the Internet, the coverage of emotional words in some specific fields such as business, economics, film and television is insufficient, and the problem of difficulty in building emotional thesaurus in new fields has become increasingly prominent [12]. To solve these problems, this study puts forward a multi-pattern SL. There are two innovative points in this study. First, semi-supervised learning (SSL) is adopted to solve the insufficient coverage of SL. The second is to adopt optimal learning to solve the difficult construction of SL in a specific field. The main structure of this study contains four parts. The first part is a summary of the research status in related fields; The second part is to construct the digital media short text analysis framework by combining various models; The third part is the effect verification analysis of the framework constructed in this study; The fourth part is the summary of this research.

**2. Related works.** As an important part of text SA, SL assigns polarity scores of text words through marking emotion scores of words, so as to judge the emotion polarity of text, which has extremely important significance for text SA. Zhou et al. introduced emotion information into human-computer dialogue and proposed a dynamic SL combining the dual replication mechanism, which was superior to other alternative methods in many aspects [20]. Wei et al. proposed a dynamic acoustic SL constructed from the acoustic lexical levels of different emotional categories to improve the ability of speech emotion recognition [17]. Navarrete et al. redesigned and constructed a SL, focusing on raising the SA in the text, enriching the emotional intensity of

words, and helping to solve the problems of cyberbullying and violence in the digital society [13]. Shoujian et al. integrated existing emotion, degree, negative and network words to achieve effective SA on Weibo, and proposed a new method to construct SL, which improved the accuracy and recall rate of the method [14]. Mahadzir et al. solved the lack of emotion words in SA research in the context of Malaysian language by using a new polarity score allocation technique to construct a new Malay-English bilingual SL, ultimately improving the comprehensive performance of mixed language SL [11]. Zha et al. designed and constructed a large SL of Chinese ultra-short comments to solve the lack of large-scale and high-precision SL in Chinese book reviews. This construction method solved the issues caused by immature segmentation techniques and imperfect language models [19]. Lijo et al. put forward a rapid polarity detection method based on multiple lexical features, and built a SL with high expansivity and high polarity detection efficiency [9]. Garg et al. proposed an emotion classification method based on centrality, and created a SL named HindiEmotionNet for Hindi, achieving satisfactory results in emotion classification and recognition [7].

Text SA refers to the analysis of text information with emotion, so as to extract and summarize the emotional tendency in it. Short text SA is mainly aimed at the subjective short text on network digital media, such as Weibo, e-commerce platforms, etc., for emotion analysis and classification. Alwehaibi et al. used three deep learning topological models and proposed an emotion classification method for short texts in Arabic dialects based on deep learning to solve the difficulties caused by the excessively complex morphology and grammar of Arabic [2]. Kota et al. combined convolutional neural networks (CNNs), bidirectional long- and short-term memory and attention mechanisms to apply deep learning to SA, achieving remarkable results and excellent performance [8]. Barnes et al. proposed a multi-tasking method to incorporate negative information into SA and optimize the priority polarity of short texts [3]. Feng et al. put forward a multi-channel CNNSA model based on multi-attention mechanism in view of the relatively limited text features of short texts, which has higher classification accuracy and lower training time cost [6]. To solve the bullying and hate speech in the Internet, Chen et al. put forward a text classification model based on CNN, which can well handle short text tasks [5]. Luo et al. proposed a text SA method integrated with neural networks for the instability of single emotion classification model in classification, which significantly improved the accuracy of text SA and effectively predicted the emotional polarity of text [10]. To solve the problem that time series in emotion is not taken into account in traditional text emotion analysis, Zhao et al. proposed the method of fusion of window word vector and classifier in the decision-making level of short text emotion analysis, which had better performance [19].

From the above content that text SA, as one of the most active research fields in natural language processing, has an extremely broad prospect in information retrieval, data mining, intelligent recommendation and other aspects. With the development of the Internet, text SA plays a more prominent role. Text SA can reflect users' subjective views through text analysis in digital media. As the basic part of it, SL plays a crucial role. The initial SL is obtained by manual summary through experience. Although it will cost labor and time, its accuracy rate is high. However, there is often a problem of insufficient coverage of emotion words in common SL, and it is difficult to construct SL in some specific fields. To solve the above problems, this study proposed to use SSL and optimized learning mechanisms to construct multi-pattern SL and build short text SA model in digital media, which has strong positive significance in emotion analysis.

**3. Digital Media Short Text SA Framework Based on Multi-pattern SL.** As the speed growth of the Internet, digital media has quickly changed the way people live and work. As the continuous growth of the Internet, new words appear constantly, which makes the coverage of the traditional emotional thesaurus insufficient, and also makes the construction of SL in a specific field difficult. Therefore, this study adopts the idea of SSL and optimal learning to improve the SL, and then applies it to the short text emotion analysis.

**3.1. A Digital Media Short Text SA Framework Based on Semi-supervised Learning SI.** DUTIR is the largest Chinese SL available, among which, How Net SL and Dalian University of Technology Chinese SL ontology library (DUTIR) are common Chinese SL. However, with the development of the Internet, many new words have emerged. This has resulted in poor coverage of DUTIR sentiment words. Therefore, this research applies the idea of SSL, extracts potential emotion words from the existing corpus of text data, identifies them, and expands DUTIR. The SA framework for short texts of digital media based on SL is shown in Figure 3.1.



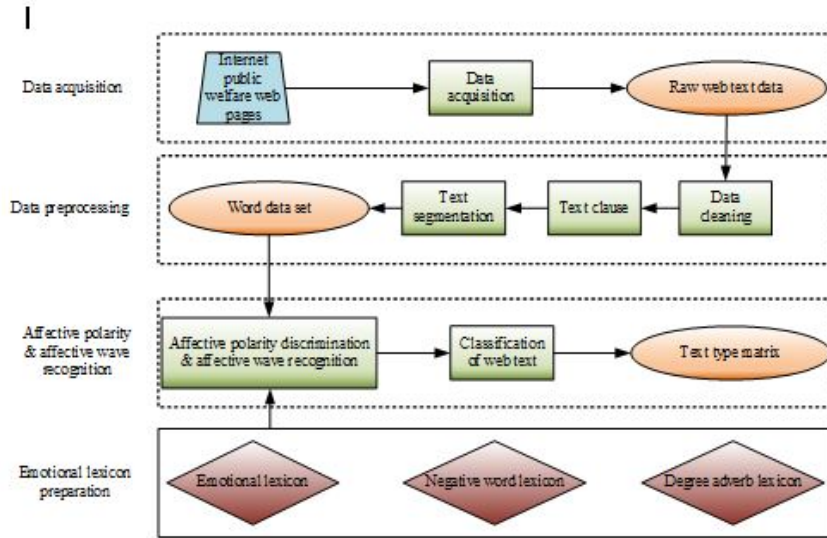


Fig. 3.1: Text SA Framework Based on SL

Emotion polarity recognition is the most important part and its function is to calculate the emotion tendency of text. Text emotion recognition calculates the emotion score of each sentence in the text and then adds all the emotion score together to get the emotion score of the whole text. If the emotion score is greater than 0, it is regarded as positive emotion; if the emotion score is less than 0, it is regarded as negative emotion; if it is equal to 0, it is regarded as neutral emotion. Its calculation is shown in Equation 3.1.

$$V = \sum_{i=1}^n v_i \quad (3.1)$$

In Equation 3.1,  $V$  denotes the emotion score of the whole text;  $n$  indicates the number of sentences in the text;  $v_i$  denotes the emotion score value of each clause. By matching the emotion lexis with the text words after word segmentation, the initial score of the emotion words is finally obtained. Then, the inversion of negative words to emotion in the sentence is calculated, which is mainly obtained by statistical negative words in the sentence, as shown in Equation 3.2.

$$v_i = (-1)^k \cdot \sum_{j=1}^m w_j \cdot w_d \quad (3.2)$$

In Equation 3.2,  $v_i$  denotes the emotion score of each clause;  $K$  indicates the number of negative words in each sentence;  $W_j$  expresses the corresponding score of emotion words;  $W_d$  refers to the weight of degree adverb;  $m$  stands for the amount of emotion words in each sentence. The next step is to calculate the emotion wave of the text. The effect of emotion wave calculation is to analyze the degree of emotion change between each sentence. The variance is calculated by the emotion score of each sentence, and then the emotion score fluctuation is obtained. The calculation is shown in Equation 3.3.

$$S = \frac{1}{n} \cdot \sum_i^n (v_i - M)^2 \quad (3.3)$$

In Equation 3.3,  $S$  represents the variance;  $M$  means the mean value of the emotion score of each text. The greater the variance, the more obvious the emotion fluctuation. IG (Information Gain), as a feature selection method with obvious effect, is usually applied to measure the importance of words in the text. It is calculated for each word in the pre-processed corpus, and then a threshold is determined. According to the level of IG, words higher than the threshold are added to the candidate emotion lexical database. The calculation is shown in Equation 3.4.

$$IG(t_i) = Entropy(S) - ExpectedEntropy(S_{t_i}) \quad (3.4)$$

In Equation 3.4,  $Entropy(S)$  means the actual entropy and  $ExpectedEntropy(S_{t_i})$  denotes the desired entropy. The specific calculation of is shown  $Entropy(S)$  in Equation 3.5.

$$Entropy(S) = - \sum_{j=1}^T P(C_j) \times \log P(C_j) \quad (3.5)$$

In Equation 3.5,  $P(C_j)$  represents the probability of  $C_j$  class documents appearing and  $T$  denotes the total number of texts.  $ExpectedEntropy(S_{t_i})$  is calculated by Equation 3.6.

$$\begin{aligned} Expected\_Entropy(S_{t_i}) = & P(t_i) \times \left[ - \sum_{j=1}^T P(C_j | t_i) \times \log P(C_j | t_i) \right] \\ & + P(\bar{t}_i) \times \left[ - \sum_{j=1}^T P(C_j | \bar{t}_i) \times \log P(C_j | \bar{t}_i) \right] \end{aligned} \quad (3.6)$$

In Equation 3.6,  $P(C_j)$  represents the probability that  $C_j$  class documents appear;  $P(t_i)$  represents the probability that word documents are included;  $P(C_i | t_i)$  represents the probability that they belong to the type  $C_j$  if they contain words  $t_i$ ;  $P(\bar{t}_i)$  refers to the probability that documents do not contain words  $t_i$ ;  $P(C_j | \bar{t}_i)$  represents the probability that they belong to the type  $C_j$  if they do not contain words  $t_i$ ;  $T$  represents the total number of texts. The SSL adopted in this study is essentially the combination of supervised learning and unsupervised learning. It realizes the semi-supervised classification algorithm by combining a large number of unlabeled sample data with a small number of labeled sample data. In this study, the labelled samples  $P$  and unlabeled samples  $U$  will be defined to identify the words belonging to them from  $U$  through the SSL method. The main steps are shown in Figure 3.2.

In SSL to enhance the performance of classification and further improve it, this study uses multi-layer perceptron for classification. The input layer is the representation of the word vector of a single word. The word vector is obtained by Word2Vec training, and its vector dimension is 200. Finally, the probability of a single word belonging to class +1 is calculated by Sigmod activation function. The structure of the multi-layer perceptron is shown in Figure 3.3.

### 3.2. The Framework of Digital Media Short Text SA Based on the Optimized Learning SL.

Digital media has changed people's way of working and living. Users can express their attitudes and views on current events, products and life at any time. The amount of data generated by digital media every day shows a blowout trend, and SA can extract users' views and opinions from it. Since there are different scenarios for SA, it means that some emotional intensity and even emotional polarity will change in different fields. Therefore, the problem of the current SL is that the universality is poor, and the general emotion word database fails to contain the domain specific emotion words. The construction of the usual domain SL is shown in Figure 3.4.

Therefore, to solve the above shortcomings, aiming at the construction of SL in a specific field, this study puts forward an optimization learning method, which adaptively adjusts the score of emotion words to build corresponding specific SL. Firstly, the emotion polarity in the text is calculated, the emotion words in the corpus are matched, and the emotion score of each emotion word is extracted in turn. After adding the emotion score of all emotion words in the text, the emotion polarity of the text is judged according to the total score. The

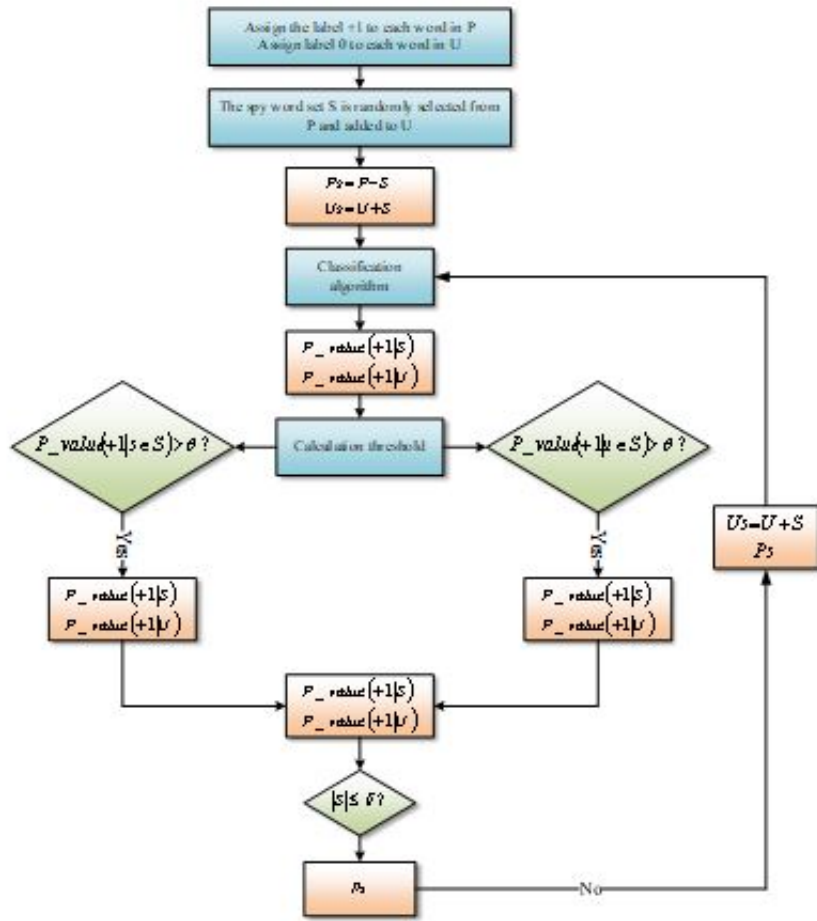


Fig. 3.2: Emotion Words Recognition Based on Semi-Supervised Learning

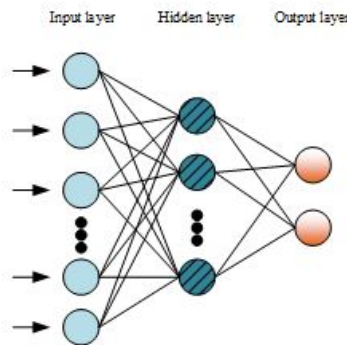


Fig. 3.3: Multilayer Augmented Perceptron

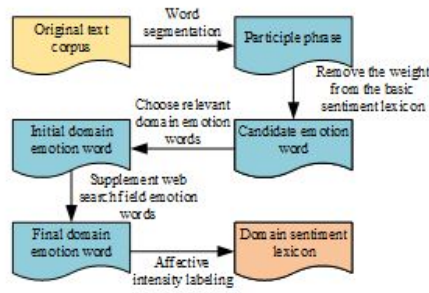


Fig. 3.4: Construction of Domain SL

calculation for the emotion polarity of the text is shown in Equation 3.7.

$$P = \sum_{w=1}^n V_w \tag{3.7}$$

In Equation 3.7,  $P$  means the emotional score of the text;  $V_w$  denotes the emotional score of the  $W_{th}$  emotional word,  $n$  and represents the total amount of emotional words in the text. If  $p > 0$ , it indicates that the text is a positive text; If  $P < 0$ , it means the text is negative. Then, it is necessary to determine the accuracy rate of emotion classification. By comparing the artificial label with the calculated emotional polarity of the text, the accuracy rate of emotion classification of the text by emotion lexicon can be calculated, as shown in Equation 3.8.

$$Accuracy = \frac{C}{N} \tag{3.8}$$

In Equation 3.8,  $C$  means the number of texts whose emotional labels are consistent with those manually marked in the emotional lexicon.  $N$  is the amount of all texts in the corpus. Finally, the score of emotion words in the SL is adjusted in a random way to improve the accuracy of emotion classification, and the effectiveness of the adjustment scheme is verified by verifying whether the accuracy of classification is increased. The optimization learning flow chart in this study is shown in Figure 3.5.

The fitness value is designed according to the accuracy of text classification by SL. For all the texts in the corpus, if the SL correctly classifies a text, a reward value 1 is added to the fitness value of the SL. If no text is correctly classified, a penalty value is subtracted from that. Its specific fitness calculation expression is shown in Equation 3.9.

$$Fitness(k, D) = \sum_{i=1}^n R(k, D, T_i) \tag{3.9}$$

Equation 3.9 shows the fitness calculation of individual  $K$  in training set  $D$  and represents  $R(K, D, T)$  the accurate measurement value of classification when individual  $K$  is used to classify the  $i$ th text  $T_i$  in training set  $D$ . Whether the text has been correctly classified can be determined by Equation 3.7. The specific classification accuracy measurement value calculation is shown in Equation 3.10.

$$R(k, D, T) = \begin{cases} 1, & \text{if Accurately Predicted} \\ -\frac{|P|}{\omega}, & \text{if not Accurately Predicted} \end{cases} \tag{3.10}$$

In Equation 3.10 shows the classification by individual, where  $K$  denotes the training set, represents the  $t$ th article text, and means the penalty parameter. In traditional genetic algorithms, there are usually problems of

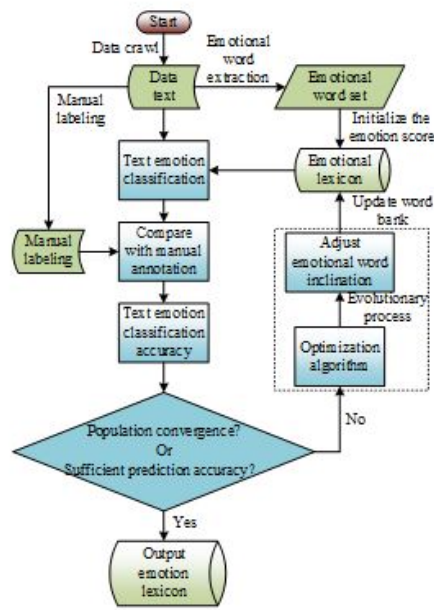


Fig. 3.5: Optimization Learning

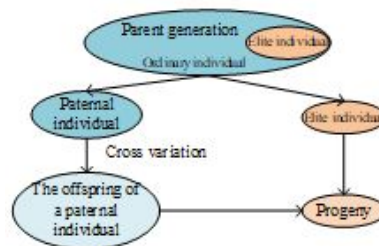


Fig. 3.6: Optimal Strategy

poor learning efficiency and poor training effect. Therefore, in order to further improve the process of emotional thesaurus construction, the optimization strategy is introduced to improve the efficiency of optimal learning. Two methods of combining optimal strategies are adopted. The first method is to directly copy the excellent individuals of the previous generation into the next generation; the second method is to select some of the better individuals as the elite set and then use the individuals and ordinary individuals to generate new individuals through cross mutation. The optimal preservation strategy designed in this study is shown in Figure 3.6.

In the text emotion analysis, the emotion polarity of the text is closely related to the polarity of the emotion words in the text. By analyzing the use of emotion words, the emotion expressed by the text can be judged. Generally speaking, if a certain emotion word often appears in positive text with a high probability, then the emotion word has a high probability of being positive emotion word; Otherwise, it is a negative emotion word. Therefore, to better guide the evolution of the score of emotion words and make the emotional polarity of text more closely related to that of emotion words, a new variation strategy is designed based on Sigmoid function

[18]. The Sigmoid function is shown in Equation 3.11.

$$S(x) = \frac{1}{1 + e^{-x}} \quad (3.11)$$

Sigmoid function is often used in the output of neurons in the hidden layer, and its value range is  $(0, 1)$ . It can map a real number to the interval  $(0, 1)$ , so as to carry out binary classification. It has the advantages of smooth and easy derivation. Equation 3.12 can be obtained by solving the inverse function of Equation 3.10.

$$f(x) = -\ln\left(\frac{1}{x} - 1\right) \quad (3.12)$$

Let the emotion word be, the data set be, and the probability of the emotion word appearing in the positive text of the data set be, plus the range of the emotion score in this paper be. Therefore, based on Equation 3.12, the variation guidance function is constructed, as shown in Equation 3.13.

$$f(P_w) = -\frac{5}{3} \times \ln\left(\frac{1}{P_w} - 1\right), P_w \in [0, 1] \quad (3.13)$$

Finally, combined with Equation 3.12, a new probability variation strategy is constructed, as shown in Equation 3.14.

$$V_w = R + f(P_w) \quad (3.14)$$

In Equation 3.14,  $R$  represents the random number,  $V_w$  indicates the emotion score value of the emotion word after variation. Equation 3.15 is the restriction of Equation 3.14.

$$\begin{cases} R \in [-10, 10] \\ V_w \in [-10, 10] \end{cases} \quad (3.15)$$

Equation 3.15 limits  $R$  value range to  $[-10, 10]$ . In this formula, if the value of  $V_w$  exceeds  $[-10, 10]$ , then correspondingly, it takes the corresponding boundary value -10 or 10.

**4. Evaluation of Short Text SA Framework of Digital Media Based on Multi-Pattern SL.** To verify the validity of the framework constructed in this study, eight widely used Chinese SA data sets were selected for experiments. The four data sets were microblog SA data sets released by NLPC, which were typical short text data sets of digital media. The other four data sets were Sina News RSS subscription channel data set (THUC News), National news data (Sogou CA), Sohu news data (Sougou CS) and a food delivery platform data (Waimai\_10k). The details of the number of positive and negative text in the experimental data set are shown in Table 4.1.

In the experiment, 80% of the data set was randomly selected to reconstruct the positive and negative sample sets, and the remaining 20 samples were utilized as test samples. To compare the effectiveness of the improved multi-pattern SL in this study, the following three scenarios were selected to compare with the framework constructed in this study: In scenario 1, the DUTIR was selected, where the sum of scores of emotion words was used as the classification criteria; In scenario 2, DUTIR and network word database (NWD) were selected, and the sum of scores of emotion words was used as the classification criteria; In scenario 3, DUTIR, NWD, and emoticon database (ED) were chosen, and the sum of values of emotion words are used as the classification standard. Firstly, data sets were used for training, and SA frameworks using different lexicon were compared. The results are shown in Figure 4.1. Where, the improved SA framework based on multi-pattern SL in this study could reach the best state after 87 training sessions, which was 65, 39 and 13 times less than scenarios 1, 2 and 3, respectively. This indicated that the convergence of the framework established in this study was better than that of other frameworks using the unimproved lexicon.

Using the selected data set, the four frameworks were tested and the accuracy rate and recall rate were obtained, as shown in Figure 4.2. Where, under the condition that the value of parameter K ranged from 1

Table 4.1: Number of Positive and Negative Comments in the NLPCC Data Set

	NLPCC2013	NLPCC2014	NLPCC2018	NLPCC2020
Positive comments	2886	2460	1049	1214
Negative comments	2047	2728	851	1158
Total comments	4960	5188	1900	2372
	THUCNews	SogouCA	SougouCS	Waimai_10k
Positive comments	6715	8977	9457	4688
Negative comments	8946	7655	7546	8154
Total comments	15661	16632	17003	12842



Fig. 4.1: Convergence of Four Situations

to 10, with the increase of parameter  $K$ , the recall rate of the SA framework established in this study based on multi-pattern sentiment lexical database increased significantly compared with other frameworks, and the value of recall reached 0.912. It was 0.102, 0.154 and 0.218 higher than scenarios 1, 2 and 3, respectively, while the accuracy rate maintained a slower fading speed compared with other models.

After each frame was trained to the best state, the test data was input for testing. Each frame as tested for 10 times. To avoid errors, the average value was taken as the final result, and the F1 value and AUC value of each model were compared. As shown in Figure 4.3, the value of the digital media short text SA framework established in this study based on mutil-pattern SL reached 0.753, which was higher than that of scenarios 1, 2 and 3 by 0.091, 0.052 and 0.032 respectively. Its AUC value also had the best performance. The results showed that in digital media short text SA, the improved lexicon proposed in this study had the best performance.

In NLPCC and the other four data sets, the recognition accuracy of the framework established in this study and scenarios 1, 2 and 3 was compared, and the average accuracy results were shown in Figure 4.4. Where, the SA framework established in this study based on the mutil-pattern SL had an outstanding performance in the direction of accuracy. And its average accuracy in NLPCC2013, NLPCC2014, NLPCC2018 and NLPCC2020 was 85.88%, 77.28%, 80.26% and 90.15%, respectively. Its overall average accuracy was 83.39%, which was 16.85%, 11.57%, and 6.72% higher than Scenarios 1, 2, and 3, respectively. Table 4.2 shows the average accuracy, positive text accuracy and negative text accuracy of NLPCC data set in this experiment.

From Table 4.2, the overall accuracy of scenario 1 on NLPCC2013, NLPCC2014, NLPCC2018, NLPCC2020 was 68.83%, 60.73%, 66.25% and 70.36%, respectively. After the combination of the network lexical database, the accuracy of scenario 2 in NLPCC2013, NLPCC2014, NLPCC2018 and NLPCC2020 was 75.95%, 64.80%, 68.94% and 77.59%, which was indeed improved compared with scenario 1. When the emoticon lexicon was added into scenario 3, its accuracy was 78.85%, 72.49%, 72.53% and 82.82%, which was better than scenario



Fig. 4.2: Precision and Recall of Four Situations

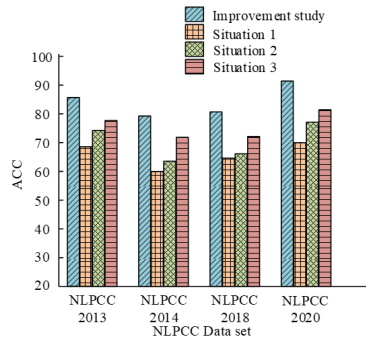


Fig. 4.3: F1 and AUC of Four Situations

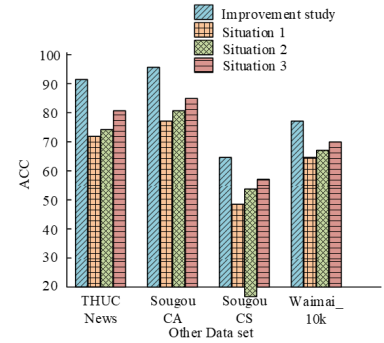
1 and scenario 2. It was suggested that network lexicon and emoticons should be considered in short text SA of digital media short text. The average accuracy of the improved SA framework based on multi-pattern SL in NLPCC2013, NLPCC2014, NLPCC2018 and NLPCC2020 was 85.88%, 77.28%, 80.26% and 90.15%, respectively, which was ahead of the three scenarios. Its recognition accuracy in positive text was 87.49%, 79.08%, 81.94% and 92.76%, which were all higher than the other three scenarios. Its recognition accuracy in negative text was 84.27%, 75.48%, 78.58% and 87.54%, which was also the highest. To sum up, after comparison with all other schemes, it showed that the framework of digital media short text SA based on multi-pattern SL proposed in this study had the best performance, and it was also the best in the classification of positive and negative texts.

**5. Conclusion.** In the field of text emotion recognition, there were problems such as insufficient coverage of emotion words and difficult construction. Therefore, a multi-pattern SL was proposed. SSL mode was adopted to solve the insufficient coverage of emotion words, and optimization learning was combined to solve the difficulty of constructing SL. The experimental results showed that the improved framework in this study needed 87 iterations to achieve the target accuracy and the optimal Loss value, which was 65, 39 and 13 times less than scenarios 1, 2 and 3, respectively. The Recall value of the improved frame in this study reached 0.912, which was 0.102, 0.154 and 0.218 higher than the other three scenarios, respectively. Its F1 value reached 0.753, which was higher than the three scenarios by 0.091, 0.052 and 0.032, respectively. The average accuracy of the four data sets was 85.88%, 77.28%, 80.26% and 90.15% respectively, and the average accuracy was 83.39%, which was 16.85%, 11.57% and 6.72% higher than the three scenarios respectively. The accuracy of positive text





(a) NLPCC Dataset ACC



(b) Other Dataset ACC

Fig. 4.4: Accuracy Mean of Four Situations

Table 4.2: Accuracy on the NLPCC Data Set

Data Set	Accuracy (%)	Average accuracy	Positive comments
NLPCC2013	-	-	-
Improvement study	85.88	87.49	84.27
Situation 1	68.83	70.12	67.54
Situation 2	75.95	75.41	76.49
Situation 3	78.85	79.24	78.45
NLPCC2014	-	-	-
Improvement study	77.28	79.08	75.48
Situation 1	60.73	61.79	59.67
Situation 2	64.80	65.87	63.72
Situation 3	72.49	73.57	71.42
NLPCC2018	-	-	-
Improvement study	80.26	81.94	78.58
Situation 1	66.25	67.25	65.24
Situation 2	68.94	69.46	68.41
Situation 3	72.53	71.48	70.57
NLPCC2020	-	-	-
Improvement study	90.15	92.76	87.54
Situation 1	70.36	71.24	69.47
Situation 2	77.59	79.86	75.31
Situation 3	82.82	84.21	81.42

in several data sets was 87.49%, 79.08%, 81.94% and 92.76%, respectively, which showed the best performance. Its accuracy in negative text was 84.27%, 75.48%, 78.58% and 87.54%, respectively, which showed the best performance. The above results showed that the framework of digital media short text SA based on multi-pattern SL established in this study could carry out digital media short text analysis accurately and efficiently, and had positive significance for the development of text SA field. However, in this study, the optimization of emotional words themselves is not directly considered, and the genetic algorithm used in this study will

inevitably slow down with the increase of time and the increase of data volume, and subsequent studies can consider changing better algorithms and further optimizing emotional words.

## REFERENCES

- [1] S. VELICHETY, S. RAM, *Finding a Needle in the Haystack, Recommending Online Communities on Social Media Platforms Using Network and Design Science*, Journal Of The Association Of Information Systems, 22 (2021) pp. 1285-1310
- [2] A. ALWEHAIBI, M. BIKDASH, M. ALBOGMI, AND K. A. ROY, *Study of the performance of embedding methods for arabic short-text sentiment analysis using deep learning approaches*, Journal of King Saud University - Computer and Information Sciences, 34, p. 201.
- [3] J. BARNES, E. VELLDAL, AND L. VRELID, *Improving sentiment analysis with multi-task learning of negation*, Natural Language Engineering, 27 (2020), pp. 249–269.
- [4] K. CHAN AND U. M. D. MEDIA CONSUMPTION: USING METRICS, *patterns and dashboards to enhance data-driven decision-making*, Journal of Consumer Behaviour, 21 (2022), pp. 80–91.
- [5] K. C. CHEN, *Verbal aggression detection on twitter comments: convolutional neural network for short-text sentiment analysis*, Neural computing & applications, 32 (2020), pp. 10809–10818.
- [6] Y. FENG AND Y. CHENG, *Short text sentiment analysis based on multi-channel cnn with multi-head attention mechanism*, IEEE Access, 9 (2021), pp. 19854–19863.
- [7] K. GARG AND A. LOBIYAL D. K. HINDI EMOTIONNET: *Scalable emotion lexicon for sentiment classification of hindi text*, ACM transactions on Asian and low-resource language information processing, 19 (2020), pp. 2–35.
- [8] V. R. KOTA AND S. D. MUNISAMY, *High accuracy offering attention mechanisms based deep learning approach using cnn/blstm for sentiment analysis*, International Journal of Intelligent Computing and Cybernetics, 15 (2022), pp. 61–74.
- [9] V. P. LIJO AND H. SEETHA, *Tweets sentiment analysis using multi-lexicon features and smo*, International Journal of Embedded Systems, 14 (2021), pp. 476–485.
- [10] S. LUO, Y. GU, X. YAO, AND W. FAN, *Research on text sentiment analysis based on neural network and ensemble learning*, Rev. D'Intelligence Artif, 35 (2021), pp. 63–70.
- [11] N. H. MAHADZIR, M. F. OMAR, M. NAWI, A. A. SALAMEH, K. C. HUSSIN, AND S. A. MELEX, *The construction of malay-english sentiment lexicon*, Computers, Materials, & Continua, 4 (2022), pp. 1789–1805.
- [12] M. NAIK, D. VASUMATHI, AND A. P. A. KUMAR, *Novel approach for extraction of distinguishing emotions for semantic granularity level sentiment analysis in multilingual context*, Recent advances in computer science and communications, 15 (2022), pp. 77–87.
- [13] A. S. NAVARRETE, C. MARTINEZ-ARANEDA, C. VIDAL-CASTRO, AND C. A. RUBIO-MANZANO, *novel approach to the creation of a labelling lexicon for improving emotion analysis in text*, The Electronic Library, 39, p. 201.
- [14] Y. U. SHOUJIAN, B. WANG, AND L. U. TING, *Sentiment lexicon construction based on improved left-right entropy algorithm*, Journal of Donghua University, 39 (2022), pp. 65–71.
- [15] J. TANG, Y. XUE, Z. WANG, S. HU, T. GONG, Y. CHEN, H. ZHAO, AND L. XIAO, *Bayesian estimation-based sentiment word embedding model for sentiment analysis*, CAAI Transactions on Intelligence Technology, 7 (2022), pp. 144–155.
- [16] Z. WANG, P. GAO, AND X. CHU, *Sentiment analysis from customer-generated online videos on product review using topic modeling and multi-attention blstm*, Advanced Engineering Informatics, 52 (2022), pp. 1588–10159.
- [17] W. WEI, X. CAO, H. LI, L. SHEN, AND W. PA., *Improving speech emotion recognition based on acoustic words emotion dictionary*, Natural Language Engineering, 27 (2020), pp. 747–761.
- [18] M. YANG, *Research on vehicle automatic driving target perception technology based on improved msrpn algorithm*, Journal of Computational and Cognitive Engineering, 1 (2022), pp. 147–151.
- [19] M. ZHA, C. HU, AND Y. SHI, *Sentiment lexicon construction for chinese book reviews based on ultrashort reviews*, The Electronic Library: The International Journal for Minicomputer, Microcomputer, and Software Applications in Libraries, 40 (2022), pp. 221–236.
- [20] Q. ZHOU, D. JI, Y. REN, T. H. D. COPYING MECHANISM, AND D. EMOTION, *Dictionary for generating emotional responses*, Neurocomputing, 454 (2021), pp. 303–312.

*Edited by:* Rajni Mohana

*Special Issue on:* Sentiment Analysis and Affective computing in Multimedia Data on Social Network

*Received:* Jun 26, 2023

*Accepted:* Jul 24, 2023



## ASPECT-BASED TEXT CLASSIFICATION FOR SENTIMENTAL ANALYSIS USING ATTENTION MECHANISM WITH RU-BILSTM

**Abstract.** Sentiment analysis has gained increasing attention from an educational and social perspective with the huge expansion of user interactions due to the Web’s significant improvement. The connection between an opinion target’s polarity scores and other aspects of the content is defined by aspect-based sentiment analysis. Identifying aspects and determining their different polarities is quite complicated because they are frequently implicit. To overcome these difficulties, efficient hybrid methods are used in aspect-based text classification in sentiment analysis. The existing process evaluates the aspects of polarity by using a Convolutional neural network, and it does not work with Big data. In this work, aspect-based text classification and attention mechanisms are used to assist in filtering out irrelevant information and quickly locating the essential features in big data. Initially, the data is collected, and then the data is preprocessed by using Tokenization, Stop word removal, Stemming, and Lemmatization. After preprocessing, the features are vectorized and extracted using Bag-of-Words and TF-IDF. Then, the extracted features are given into word embeddings by GloVe and Word2vec. It uses Deep Recurrent based Bidirectional Long Short Term Memory (RU-BiLSTM) for aspect-based sentiment analysis. The RU-Bi-LSTM method integrates aspect-based embeddings and an attention mechanism for text classification. The attention mechanism focuses on more crucial aspects and the bidirectional LSTM to maintain context in both ways. Finally, the binary and ternary classification outcomes are obtained using the final dense softmax output layer. The proposed RU-BiLSTM uses four reviews and two Twitter datasets. The results of the studies demonstrate the efficacy of the RU-BiLSTM model, which outperformed aspect-based classifications on lengthy reviews and short tweets in terms of evaluation.

**Key words:** aspect-based sentiment analysis, attention mechanism, aspect level, word2vec, TF-IDF, Convolutional Neural Network, big data, BiLSTM

**1. Introduction.** The Sentiment Analysis is useful in finding the feelings expressed in the text’s sections. Aspect-Based Sentiment Analysis (ABSA), a pioneering technique, asserts that there are three stages at which SA research is carried out: text, sentence, word, or aspect. SA implies that each document expresses a view about a single entity at the level of the document. It is supposed that the text only covers one subject, yet this is frequently untrue. Every sentence is viewed as a separate entity in a sentencing-level strategy, and it is thought that the sentence should only include one assumption. Two tasks of the sentence-level analysis are grouping estimation and subjective characterization.

The sentiment analysis (SA) of texts aims to gather and examine facts from the private data placed on social media. Nowadays, there is an interest in the area of natural language processing (NLP) related to SA because of its broad range of educational and commercial uses and the expansion of social networks. As a result, various tools and techniques have been recommended to determine the polarity of texts. In most SA applications, polarity recognition is a crucial binary classification problem. Earlier SA methods relied on deep procedures and cleverly engineered effective characteristics to achieve acceptable polarity classification outcomes [11].

To overcome the earlier restrictions, learning word embedding has been recommended in several recent kinds of research [7, 28, 30]. A thick real-valued vector called word embedding, which considers different lexical relationships, is produced by a neural language model [31, 33]. As a result, deep neural networks (DNN) frequently use word embedding as their input in existing NLP works. Experts from a wide range of fields, including computer vision [32], finance [3], medical informatics, and multimedia sentiment analysis [2], have been paying more and more attention to DNNs in recent years.

As e-commerce grows, increasing numbers of individuals are eager to express their views and thoughts on branded items after using them, resulting in an enormous amount of remark messages. These little compositions frequently have a significant subjective element, occasionally incorporating many psychological characteristics within a single sentence. Short text remarks are also incredibly informal. Because of this, the text’s subject is ambiguous, difficult to locate, linguistically inconsistent, and, more concerning, challenging for scholars to use effectively. Nevertheless, most recent studies on analyzing brief remark texts’ sentiments use a coarse-grained

approach, whereby just one sentiment trend is gleaned from a sentence [9].

A simple positive signal will be given to this remark in coarse-grained sentiment analysis. In actuality, the user's feedback was merely positive regarding the product's look and unfavorable regarding its performance. As it has been shown, coarse-grained sentiment classification is unable to capture the particulars that users care about fully. The emotive inclinations of various features in some customer review ontologies can be identified using aspect-based sentiment analysis. Depending on this benefit, aspect-based sentiment classification of brief user review texts can assist businesses in enhancing specific products while also assisting consumers in making better choices.

This work aimed to use deep learning research techniques to increase the accuracy of the aspect-based Attention mechanism for sentiment analysis in response to the issues mentioned earlier. This approach can record remarks at the aspect level and extract local feature information. Furthermore, it can prevent gradient expansion and disappearing problems, enhance model fidelity, and lower computing overhead. In addition, this approach can be applied practically to the emotional evaluation of brief texts like microblog responses and social attitudes postings. Those are important in relevant domains in all spheres of life.

**1.1. Motivation.** The motivation for this research is to introduce a new method for aspect-based sentiment analysis that can efficiently handle big data and accurately identify the polarity of implicit aspects in text. The work highlights the challenges of aspect-based sentiment analysis and the limitations of existing methods, which cannot work with big data. The proposed method uses pre-processing, vectorization, and deep learning models like RU-BiLSTM with an attention mechanism to classify text based on aspects and their polarities. The research aims to demonstrate the efficacy of this new method through experiments on four review datasets and two Twitter datasets. This work provides a solution to the challenges of aspect-based sentiment analysis and highlights the importance of accurate sentiment analysis in the context of social and educational applications.

**1.2. Objectives of the study.** The following objectives are the focus of this research study:

1. This study makes use of LSTM and BiLSTM to analyze two kinds of datasets. They are long reviews and short tweets from social media.
2. By testing with both long and short user evaluations, it explores the capacity of LSTM and BiLSTM as well as the attention process in gathering contextual information.
3. We provide an attention mechanism to effectively improve the emotion polarities of words, identify important data in the text, and collect the phrase within the text that are highly connected to the extended gap and encoded dependency.

**1.3. Contributions of this study.** The main contribution of the suggested method is given below:

1. The information was first cleaned during the preprocessing stage to remove any special symbols, syntax, English, numerals, etc.
2. The features are vectorized by using Bag-of-Words and TF-IDF.
3. Word2Vec and GloVe are used for extracting features.
4. The attention mechanism effectively increases the sentiment polarity of the words, detects the important data in the text and catches the words within the text which are strongly connected to long space and encoding dependency.
5. Attention mechanism provides specific attention to the data produced from the hidden layers of RU-BiLSTM.

The remaining part of our research article is written as follows: Section 2 discusses the related work on Chinese corpus comments, Attention Mechanism, and Aspect level Sentiment Analysis. Section 3 shows the general working methodology of the proposed work. Section 4 evaluates the implementation and results of the proposed method. Section 5 concludes the work and discusses the result evaluation.

**2. Related Works.** The majority of classic SA investigations have used supervised machine learning techniques as their primary classification or clustering module [4]. These methods classified and displayed user-created texts that contained sentiment by utilizing n-gram features and bag-of-words (BOW) techniques [16]. These elements address the drawbacks of basic BOW strategies, such as failing to consider word order and syntactic structures [36]. Utilizing n-gram features is a disadvantage because the resulting feature space is

dimensional, especially when  $n \geq 3$ . Recently, feature selection strategies have been widely used to address this issue [35].

SVM, Naive Bayes (NB), and artificial neural networks (ANN) are a few of the often used techniques for extracting users' meanings from their texts and have provided promising results [19, 36, 24]. The supervised approaches have several drawbacks, including the fact that training can be time-consuming and slow at times. Numerous unsupervised lexicon-based approaches have been put forth to address these issues [15, 32]. These methods are quick, easy, and scalable. They are less efficient than their supervised counterparts, nevertheless, because they heavily rely on the lexicon [13, 1]. Lexicon-based techniques also suffer from field reliance, which limits their use in disciplines without specialized lexicons.

Aspect extraction [26] was addressed using a linear chain conditional random field after the author addressed aspect removal as a sequence classification issue. Text representations are separated from extracting features and model training in conventional methods (such as creating sentiment dictionaries), which concentrate on word embedding and extraction of features. Short texts include a high degree of randomness, confusion, inconsistency, and other qualities that make it easy for feature dispersal and contextual independence issues to arise during the processes of word embedding and extraction of features. When employing conventional sentiment analysis methods, all of these variables may result in decreased feature extraction accuracy and disconnection of contextual semantic links [34].

In SA, CNNs are used to extract specific characteristics. Those methods are helpful whenever the text is lengthy and particular local properties, such as  $n$ -grams, matter. For example, to analyze the sentiments at the data level, the researcher presented a CNN-based model that made use of optimal word embedding [12]. Their approach enhanced spatial, syntactic and semantic, and lexical features using pre-trained GloVe and Word2Vec embedding [25], but this work did not take into account the varying significance of terms and long dependency.

In the latest days, the attention mechanism was employed to improve DNN modeling optimization by letting the DNNs know where to focus their learning efforts. For example, [23] proposed a global pooling strategy architecture and one BiLSTM layer for binary sentiment categorization. The researcher [8] suggested a mixed model that incorporates Bi-LSTM, CNN, and the attention mechanism, AC-BiLSTM, for text categorization and information retrieval. On top of the word embedding layer, their models were using a one-dimensional CNN layer to automatically extract characteristics, a BiLSTM layer to extract long dependencies and an attention mechanism to concentrate on important text regions. The co-occurrence of both long and short dependency was not taken into account by the AC-BiLSTM framework.

Over the past several years, deep neural networks have become increasingly popular for solving pattern classification and machine learning challenges [14]. In the areas of speech recognition and computer vision, deep learning-based models have also demonstrated outstanding results [20, 17, 6]. Additionally, RNN-based networks have specifically had more success in the field of NLP deep networks [18, 22]. While conducting a sentiment analysis, word modeling is a crucial activity. Since the document level comprises more stored facts and is retrieved at the global level as opposed to the local level, a document-level sentiment classification is seen as a primary stage and is favored over sentence-level sentiment analysis. The amazing accomplishments in sentiment analysis that neural nets have achieved are due to their capacity to handle sequences of various lengths. Between these neural nets, the long short-term memory units are noteworthy [27, 36, 21, 24, 29]. Recently deep learning technologies with multi-layer neural networks are widely used to improve the performance of prediction and classification [10, 5].

**3. Proposed Methodology.** The proposed method for Aspect-based text classification uses the Attention Mechanism and RU-BiLSTM. Initially, the dataset is pre-processed and then the preprocessed data is given into feature extraction. Next, the extracted features are classified using RU-BiLSTM. We also suggested an attention mechanism-based LSTM and BiLSTM technique to categorize the polarity of reviews and tweets. The suggested approach aims to clarify the semantic similarities among features in addition to solving the issue of long-term dependency. The overall design of the suggested method is shown in Figure 3.1.

**3.1. Dataset Collection.** Long and short datasets were used in our study to perform sentiment analysis and aspect-based text categorization tasks. The following are the datasets' specifics:

1. APP: This collection for Android apps [10] includes 752,937 Amazon product reviews and related metadata.

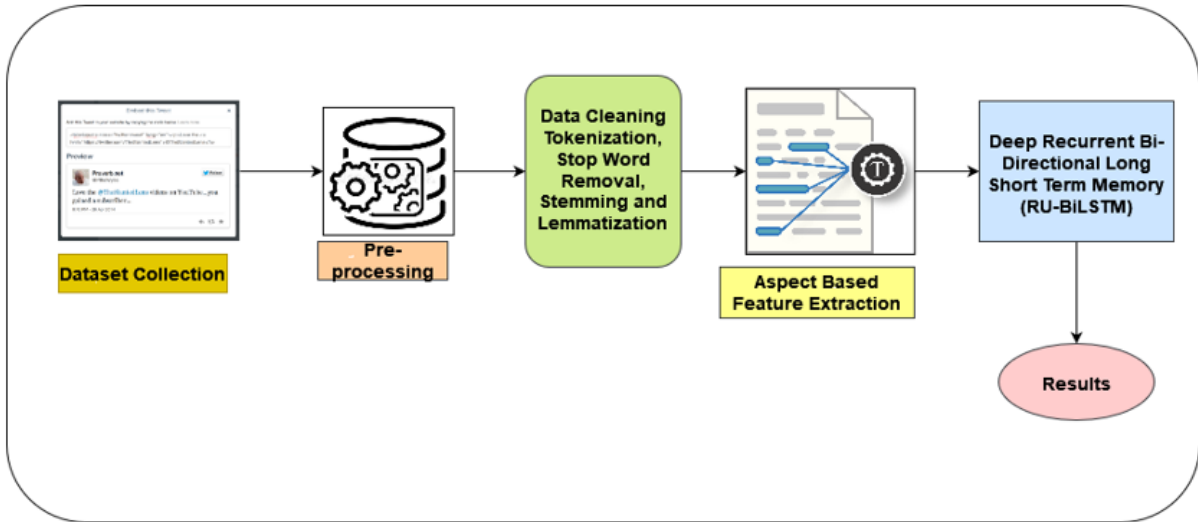


Fig. 3.1: Overall design of the Suggested Method

Table 3.1: Descriptions of Dataset [28]

Type of Dataset	Dataset Used	The total amount of Dataset	Positive	Negative
Reviews	App	752,748	133,897	112,774
	Kindle	995,619	59,548	58,157
	Electronics	1,730,148	199,865	198,912
	Cd	1,397,941	99,472	97,946
Tweets	Airline Tweets	15,649	3763	3483
	Sentiment 140	1,730,520	920,890	913,700

- Kindle: In this [10] the Kindle data contains 996,732 Amazon product reviews and related metadata. [10]
- Electronics: 1,732,458 Amazon manufacturing remarks and related meta-data are included in this dataset for electronics [10].
- CDs: The 1,097,592 products metadata and customer reviews for CDs and Vinyl [10] are from Amazon.
- Airline Twitter: Here the dataset for airlines [10] is given and 14,641 tweets regarding significant issues with U.S. airlines are included in this dataset from February 2015.
- Sentiment140: It consists of 1,730,520 tweets divided into positive and negative classes, and was created at Stanford University [10].

The parameters of the datasets utilized in the suggested model are shown in Table 3.1 along with additional information.

**3.2. Data Preprocessing.** The acquired data is cleansed during this procedure so that it is the sole data required for text classification. Preprocessing includes several steps, including review cleaning, which is useful for removing superfluous words and converting upper case letters in the text to lower case letters for text categorization, and stop words removal, which is useful for eliminating conjunction phrases as well as other special characters, like emoticons, that are not utilized for text classification, stemming, which is useful for turning all words into basic words, and tokenization, which is useful for tokenizing text.

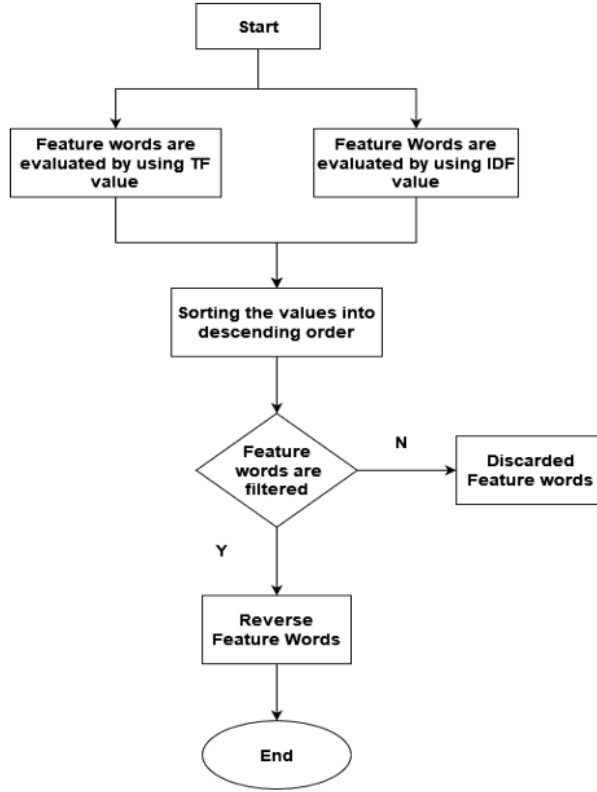


Fig. 3.2: Workflow of TF-IDF

### 3.3. Aspect-Based as Feature Extraction using Bag-of-Words and TF-IDF.

**3.3.1. TF-IDF.** An accepted feature extraction technique is the TF-IDF method. The TF-IDF extraction algorithm is much more accurate when compared to other algorithms. As a result, this work implemented vectorization analysis for reviews and tweets using the TF-IDF feature extraction technique. We are capable of determining whether another word was crucial in these textual samples by using the TF-IDF computation. The precise formula for calculation is as follows.

$$f(\omega) = TF(\omega) * IDF(\omega) = TF(\omega) * \frac{\log N}{n(\omega) + 1} \quad (3.1)$$

After calculating the  $TF(\omega)$  and  $IDF(\omega)$  values independently, the overall weight value of TF-IDF is produced and arranged in ascending order. The dimensionality reduction process involved selecting the first five phrases with the greatest values. In Figure 3.2, the procedure is displayed.

**3.3.2. TF-IDF Keyword Table.** Numerous methods for numerical operations are available in the Python Scikit-learn machine learning toolkit, which also offers the Tfidf Transformer function needed for the TF-IDF method mentioned in this study. The weight was determined using the aforementioned method to eliminate keywords appropriate keywords 90 keywords from assessments of reviews and tweets were filtered out for this article, and then their corresponding values were determined.

**3.3.3. Bag-of-Words.** A method called bag of words is being used in natural language processing to count how many times each word appears in a text or review. Any arbitrary number of words, or n-grams, can be used to describe a phrase or token. The (1, 2) n-gram frequency is used in this investigation. The framing of unigrams, digrams, and trigrams from a phrase is shown in Fig. 3.3. Due to the Bow model's consideration

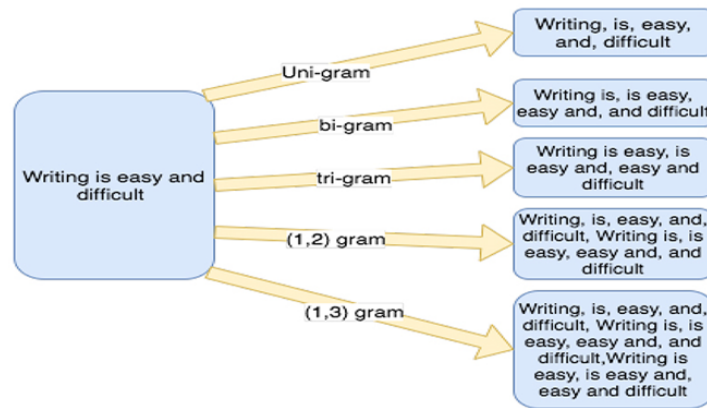


Fig. 3.3: Different types of grams used in texts

of all terms without taking into account the fact that some phrases are extremely consecutive in the corpus, a big matrix that is operationally costly to train is created.

**3.4. Word Embeddings.** It uses two types of word embeddings such as GloVe and Word2Vec. Here the sentence and aspects are embedded by the above methods.

**3.4.1. GloVe .** The input review or comment matrix  $WE_g \in R^{n \times e}$  was made using a pre-trained GloVe embedding matrix, where  $e$  and  $n$  stand for the embedding dimensions and overall word count, accordingly. The max number of characters  $w$  or the padding duration,  $t \in [1, n]$ , found in the remark for embedding purposes is repress ended  $c \in R^n$ , as seen below:

$$we_t = WE_g we_t, t \in [1, n] \tag{3.2}$$

**3.4.2. Word2Vec.** Setting low - dimensional feature-dense matrices for subsequent layers of neural networks is the main goal of embeddings. Our suggested method, RU-BiLSTM, can undertake detailed feature extraction with the use of n-grams and TF-IDF, improving classification even more. For improved text representation, a large volume of Chinese Comments text adjusted to a million terms was dragged from web reviews.

The input-padded segments are incorporated into the word embedding hidden layer before the data is fed to our main model. We offered a document  $D$  of  $M$  words, where  $D = \{x_1, x_2, \dots, x_M\}$ , and  $x_i$  stands for each phrase in  $D$ . For each word in  $D$ , we used the embedded matrix  $W^k$  as a dictionary and lookup table, where  $W^k \in R^{d^w \times |V|}$ ; in this case,  $V$  is our language dimension, which is set to 10,000, and  $d^w$  is our embedding size. While the user must supply the hyperparameter  $d^w$ , the weight vector  $W^k$  must be trained. This weight matrix's initialization is random. Word embedding's goal is to transform  $x_i$  into  $p_i$ , which rs essentially the result of the weight matrices and vectors.

$$p_i = W^k \cdot v^i \tag{3.3}$$

Here  $v^i$  is the vector of vocabulary size.

**3.5. RU-BiLSTM.** A more in-depth look at our model is also shown in Figure 3.4, where it is clear that BiLSTM encoding is first used at the aspect level before being used at the sentence level. Furthermore, both aspect and sentence levels employ an attentiveness strategy. The layers of our model RU-BiLSTM will be discussed below.

**3.5.1. Recurrent Neural Network.** Both the traditional LSTM and its variation BiLSTM were empirically assessed. RNN can record historical data. In other words, while basic feed-forward networks just ahead



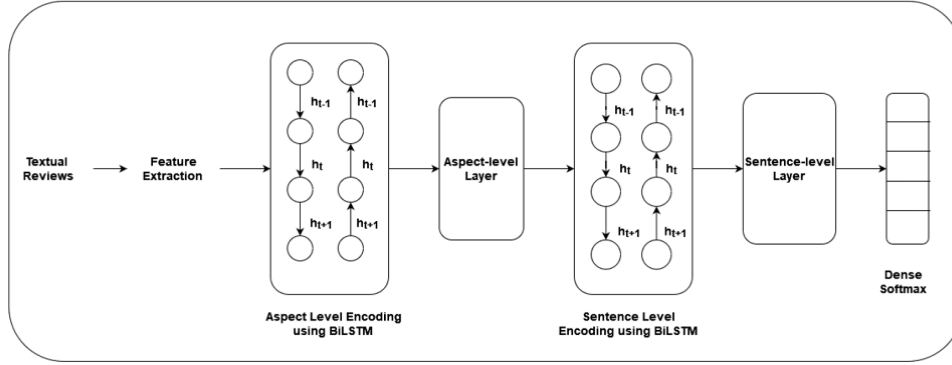


Fig. 3.4: Structure of Bi-directional Long Short Term Memory

the incoming input and lack memory, RNNs can maintain the prior outputs. Simple RNNs cannot recall long time stamps; they can only recall the details of small time stamps. Long-term dependency is the issue, and LSTM can be used to solve it. Modern LSTM devices can address this issue of long-term dependence. In addition to preserving context data, the recurrent design of LSTM can also solve the gradient descent's "blowing up and disappearing" problem.

The LSTM's gated components provide it the ability to manage the flow and select what to update and what to neglect. A recurrent module of single layer with a tanh squelching function makes up the simple RNN. The basic RNN can be theoretically described as follows as long as, input neurons as  $x_t$ , a hidden state in output is  $h_t$ , and the preceding hidden state  $h_{t-1}$  as output.

$$h^t = g_h(W_i x^t + W_R h^{(t-1)} + b_h) \quad (3.4)$$

$$y^t = g_y(W_y h^t + b_y) \quad (3.5)$$

Here, weighted matrix is  $W$ , bias as  $b$ , hidden output as  $h^t$ , squashing function as GH, and it is the outcome.

Cell modes that function as conveyors are the main concept underpinning LSTM. This data is transmitted along the belt conveyor in a series of cycles, passing across various cell states with a minimal amount of linear interactions. Using gates, some data is added to or subtracted from the cell states throughout this journey. The traditional LSTM comprises four system memories, each with three multiplicative units called gates and a cell state  $c$ . As seen, the gates are input gate  $i(t)$ , output gate  $o(t)$ , and forget gate  $f_t$ .

$$f^t = \sigma(W_{xf}x_t + W_{hf}h^{t-1} + b_f) \quad (3.6)$$

$$i^t = \sigma(W_{xi}x_t + W_{hi}h^{t-1} + b_i) \quad (3.7)$$

$$c_t^{\sim} = \tanh(W_{xc}x_t + W_{hc}h^{t-1} + b_c) \quad (3.8)$$

The cell state being updated

$$c_t = f_t \otimes c_{t-1} + i_t \otimes c_t^{\sim} \quad (3.9)$$

$$o_t = \sigma(W_{xo}x_t + W_{ho}h^{t-1} + b_o) \quad (3.10)$$

$$h_t = o_t \otimes \tanh(c_t) \quad (3.11)$$

Thus, the sigmoid function is denoted by the symbol, while the forget gate, input gate, cell state, output gate, and hidden state are denoted by the symbols  $f_t, i_t, c_t, o_t,$  and  $h_t,$  correspondingly. In comparison,  $W$  indicates that  $b$  is the bias and that the weight matrix relates to every layer. Finding information that can be removed from the cell state and is not required is the first stage in LSTM. This decision is made by the sigmoid layer, also referred to as the forgotten layer.

The input gate layer, a sigmoid layer that determines the numbers to be modified, and the layer, that creates an angle of potential possible numbers to be placed in cell state  $c \otimes t,$  are two-pass layers that we use to decide whether additional knowledge is to be supplied. As noted in Equation 3.9, the new state of cell  $ct$  is then merged with the previous state of cell  $ct_1.$  The final step is to run a sigmoid layer between 0 and 1 to decide which parts are included in the cell state as the output of the algorithm, which is then multiplied based on element-wise (represented by the symbol) with a layer tanh that takes the cell state as a dispute. The function ‘tanh’ limits the range of numbers to those between 0 and 1. As a result, concealed output  $h_t.$

**3.5.2. Attention mechanism with Bidirectional Network.** One shortcoming of the traditional LSTM is that it only takes into account past outputs. We may think about the environment in both ways and utilize the past and the future contexts thanks to our Bi-LSTM driven by the attention mechanism. This is accomplished by using two distinct hidden bidirectional layers that are later merged into a single layer output and sent into the attention layer. The forward and backward layers of our bidirectional network, which are designated as  $L_f$  and  $L_b,$  accordingly, are hidden LSTM layers. The two layers move back and forth through the SC feature sequences from 1 to n. Equations 3.12 and 3.13 provide a mathematical representation of the network.

$$\vec{h}_f = \vec{L}_f(SC)_n, \quad n \in [1, 100] \quad (3.12)$$

$$\overleftarrow{h}_b = \overleftarrow{L}_b(SC)_n, \quad n \in [1, 100] \quad (3.13)$$

The BiLSTM layer decides how words are annotated in both directions before summarizing the data. The attention mechanism enables one to focus on a phrase that is more significant to a particular class of sentiment content. Key features are given more weight by the attention layer, whereas unimportant features are ignored. The thick softmax layer continues to process the gathered characteristics before producing the final result.

*Attention Mechanism.* Every output  $h_t$  is from hidden layer of LSTM layers is subjected to the attention mechanism to give words with different contributions distinct weights. One common method of giving various weights to several words in a sentence is to use a weighted mixture of all hidden states. The following is how the context vector is to be calculated using the attention model:

$$\mu_t = \tanh(W.h_t + b) \quad (3.14)$$

$$\alpha_t = \frac{\exp(\mu_t^T u_w)}{\sum_t \exp(\mu_t^T u_w)} \quad (3.15)$$

$$c = \sum_t^T \alpha_t h_t \quad (3.16)$$

Here  $u_w$  represents the vector representation which is initialized arbitrarily and eventually learned at training stage, and  $u_t$  represents a concealed version of  $h_t.$  The input signal has a total of  $T$  time steps, and each time step  $t$  has a weight calculated for every state  $h_t.$  After that, the weightings  $\alpha_t$  are calculated as given in Equation 3.14. These important weights are gradually combined into  $c$  as stated in Equation by using a weight value to them 3.16.  $c$  symbolizes the vector and combines all of the textual review’s word data.

Table 4.1: Simulation Parameters

Parameters Used	Values
Dimensions (D)	100
Hidden Bi-LSTM units	155
Total comments for reviews	100
Total comments for tweets	55
Activation Function	Relu
Optimizer	Softmax
Kernel-size	4
Convolution Layer	64
Loss function	Binary-cross entropy

**3.5.3. Softmax layer.** The vector  $V_{gap}$  output was quickly passed into the Softmax layer to perform sentiment analysis, as illustrated in the formula below:

$$\hat{y} = softmax(WV_{gap} + b) \quad (3.17)$$

The goal of cross-entropy was initiated to represent the discrepancy between the projected emotional category  $\hat{y}$  and the actual sentimental category  $y$  to assess the suggested model.

$$Loss = - \sum_i y_i \log \hat{y}_i$$

Here  $i$  stand for the phrase's index.

**4. Result Analysis.** This section shows the tests that were carried out to evaluate the RU-BiLSTM model's SA and aspect-based text classification results on several testing samples. Following a discussion of the findings is a description of the experimental design and foundational procedures.

**4.1. Simulation Parameters.** Tensorflow 1.13.1 with Keras 2.24 libraries built in Python 3.7.1 and an Ubuntu 16.04 machine with a Core Tetranuclear i7-7700k CPU and a GTX1080 Ti GAMING X 11GB GPU were used to apply the RU-BiLSTM model. The Tokenizer approach employs 100,000 words to build the input comment matrix  $C$ . We fixed the padding values to 45 and 100, accordingly, assuming that the first 45 and 100 words of remarks in the tweet and review datasets, respectively. The pre-trained, publicly accessible GloVe and word2Vec models were used in the current investigation as the embedding layer's weights. It used the "Gigaword 5 + Wikipedia 2014" version of GloVe, which has a vocabulary size of 400,000 words and six billion tokens. The embedding length of 300 was employed for the embedding layer. Table 4.1 displays further parameter selections made use of in the suggested model.

For assessing the effectiveness of the suggested approach, four assessment principles—Accuracy, Precision, Recall, and F1 measures—were used. In tasks involving SA and aspect-based text classification, these standards are frequently used. This is how these standards are calculated:

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100 \quad (4.1)$$

$$precision = \frac{TP}{TP + FP} \times 100 \quad (4.2)$$

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

Table 4.2: Experimental outcome of Kindle Dataset

Techniques Used	Classes Used	Accuracy	Precision	Recall	F1-Score
SS-BED	Positive	0.8910	0.9461	0.8308	0.8827
	Negative		0.8521	0.9514	0.8979
AC-Bi-LSTM	Positive	0.9074	0.9018	0.8553	0.9018
	Negative		0.9122	0.9595	0.9122
DNN-WHAT	Positive	0.9372	0.9612	0.9123	0.9377
	Negative		0.9193	0.9614	0.9343
RU-BiLSTM	Positive	<b>0.9585</b>	<b>0.9732</b>	<b>0.9256</b>	<b>0.9478</b>
	Negative		<b>0.9317</b>	<b>0.9775</b>	<b>0.9495</b>

Table 4.3: Experimental results of APP dataset

Techniques Used	Classes Used	Accuracy	Precision	Recall	F1-Score
SS-BED	Positive	0.8370	0.9371	0.8778	0.8127
	Negative		0.8731	0.9014	0.8639
AC-Bi-LSTM	Positive	0.9104	0.9428	0.9012	0.9074
	Negative		0.9512	0.9075	0.9174
DNN-MHAT	Positive	0.9256	0.9505	0.9006	0.9244
	Negative		0.9063	0.9523	0.9304
Proposed RU-BiLSTM	Positive	<b>0.9397</b>	<b>0.9624</b>	<b>0.9106</b>	<b>0.9318</b>
	Negative		<b>0.9210</b>	<b>0.9605</b>	<b>0.9402</b>

$$F1 - score = \frac{2 * Precision * Recall}{Precision + Recall} \quad (4.4)$$

The proposed RU-BiLSTM method is compared with existing methods such as DNN-MHAT, AC-BiLSTM and SS-BED. It uses long and short datasets for evaluation.

**4.2. Evaluation results of Long Reviews.** In this long review, four types of datasets were used such as Kindle, APP, CD, and Electronics datasets. The positive and negative aspects were classified. Tables 4.2, 4.3, 4.4, 4.5 show the experimental results of the review dataset.

The table 4.2 shows the results of evaluating four different techniques (SS-BED, AC-Bi-LSTM, DNN-WHAT, and Proposed RU-BiLSTM) on a binary classification task, where the classes are Positive and Negative. The evaluation metrics used are Accuracy, Precision, Recall, and F1-Score for both classes. The results show that all techniques performed well, with the Proposed RU-BiLSTM achieving the highest overall performance with an Accuracy of 0.9585. It also achieved the highest scores for Precision, Recall, and F1-Score for the Positive class, indicating that it was able to correctly identify the Positive class with a high level of precision and recall.

The table 4.3 presents experimental results of four different techniques used to classify classes of the APP dataset. The dataset consists of text data and the techniques used are SS-BED, AC-Bi-LSTM, DNN-MHAT, and Proposed RU-BiLSTM. The SS-BED technique achieved an accuracy of 0.8370 for the positive class and 0.8731 for the negative class. The precision for the positive class was 0.9371, and the recall was 0.8778, resulting in an F1-score of 0.8127. For the negative class, the precision was 0.9014, and the recall was 0.8639, resulting in an F1-score of 0.8370. The AC-Bi-LSTM technique achieved an accuracy of 0.9104 for the positive class and 0.9512 for the negative class. The precision for the positive class was 0.9428, and the recall was 0.9012, resulting in an F1-score of 0.9074. For the negative class, the precision was 0.9075, and the recall was 0.9174, resulting in an F1-score of 0.9124.

The DNN-MHAT technique achieved an accuracy of 0.9256 for the positive class and 0.9063 for the negative class. The precision for the positive class was 0.9505, and the recall was 0.9006, resulting in an F1-score of

Table 4.4: Experimental results of CD Dataset

Techniques Used	Classes Used	Accuracy	Precision	Recall	F1-Score
SS-BED	Positive	0.8512	0.8971	0.8568	0.8723
	Negative		0.8798	0.8914	0.8930
AC-Bi-LSTM	Positive	0.8704	0.9208	0.8712	0.8474
	Negative		0.9012	0.8975	0.8974
DNN-MHAT	Positive	0.8913	0.9188	0.8578	0.8706
	Negative		0.8692	0.9306	0.8959
Proposed RU-BiLSTM	Positive	<b>0.9071</b>	<b>0.9243</b>	<b>0.8706</b>	<b>0.9089</b>
	Negative		<b>0.8712</b>	<b>0.9434</b>	<b>0.9207</b>

Table 4.5: Experimental Results of Electronic Dataset

Techniques Used	Classes Used	Accuracy	Precision	Recall	F1-Score
SS-BED	Positive	0.8512	0.8671	0.8668	0.8833
	Negative		0.8818	0.8812	0.8607
AC-Bi-LSTM	Positive	0.8704	0.9078	0.8904	0.8704
	Negative		0.8972	0.8741	0.8520
DNN-MHAT	Positive	0.9112	0.9411	0.8777	0.9065
	Negative		0.8821	0.9482	0.9127
Proposed RU-BiLSTM	Positive	<b>0.9245</b>	<b>0.9537</b>	<b>0.8837</b>	<b>0.9189</b>
	Negative		<b>0.8977</b>	<b>0.9599</b>	<b>0.9227</b>

0.9244. For the negative class, the precision was 0.9523, and the recall was 0.9304, resulting in an F1-score of 0.9412. The Proposed RU-BiLSTM technique achieved the highest accuracy of 0.9397 for the positive class and 0.9210 for the negative class. The precision for the positive class was 0.9624, and the recall was 0.9106, resulting in an F1-score of 0.9318. For the negative class, the precision was 0.9605, and the recall was 0.9402, resulting in an F1-score of 0.9502.

The table 4.4 presents the experimental results of four different techniques used to classify classes of the CD dataset. The techniques used are SS-BED, AC-Bi-LSTM, DNN-MHAT, and Proposed RU-BiLSTM. The SS-BED technique achieved an accuracy of 0.8512 for the positive class and 0.8798 for the negative class. The precision for the positive class was 0.8971, and the recall was 0.8568, resulting in an F1-score of 0.8723. For the negative class, the precision was 0.8914, and the recall was 0.8930, resulting in an F1-score of 0.8922.

The AC-Bi-LSTM technique achieved an accuracy of 0.8704 for the positive class and 0.9012 for the negative class. The precision for the positive class was 0.9208, and the recall was 0.8712, resulting in an F1-score of 0.8474. For the negative class, the precision was 0.8975, and the recall was 0.8974, resulting in an F1-score of 0.8975. The DNN-MHAT technique achieved an accuracy of 0.8913 for the positive class and 0.8692 for the negative class. The precision for the positive class was 0.9188, and the recall was 0.8578, resulting in an F1-score of 0.8706. For the negative class, the precision was 0.9306, and the recall was 0.8959, resulting in an F1-score of 0.9130.

The Proposed RU-BiLSTM technique achieved the highest accuracy of 0.9071 for the positive class and 0.8712 for the negative class. The precision for the positive class was 0.9243, and the recall was 0.8706, resulting in an F1-score of 0.9089. For the negative class, the precision was 0.9434, and the recall was 0.9207, resulting in an F1-score of 0.9320.

The table 4.5 shows the experimental results of different techniques used on an electronic dataset, where each technique is evaluated based on its accuracy, precision, recall, and F1-score for positive and negative classes. The first technique used is SS-BED, which achieved an accuracy of 0.8512 for the positive class and 0.8818 for the negative class. Its precision is 0.8671 for the positive class and 0.8812 for the negative class,

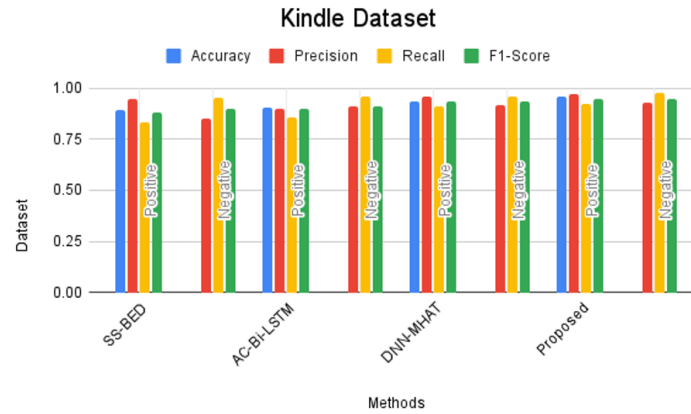


Fig. 4.1: Results of Kindle dataset

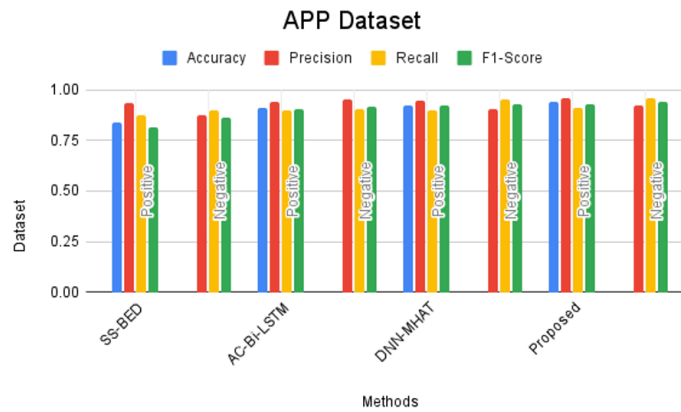


Fig. 4.2: Results of APP dataset

while its recall is 0.8668 for the positive class and 0.8607 for the negative class. Its F1-score is 0.8833 for the positive class and 0.8520 for the negative class.

The second technique is AC-Bi-LSTM, which achieved an accuracy of 0.8704 for the positive class and 0.8972 for the negative class. Its precision is 0.9078 for the positive class and 0.8741 for the negative class, while its recall is 0.8904 for the positive class and 0.8520 for the negative class. Its F1-score is 0.8704 for the positive class and 0.8520 for the negative class. The third technique is DNN-MHAT, which achieved an accuracy of 0.9112 for the positive class and 0.8821 for the negative class. Its precision is 0.9411 for the positive class and 0.9482 for the negative class, while its recall is 0.8777 for the positive class and 0.9127 for the negative class. Its F1-score is 0.9065 for the positive class and 0.9127 for the negative class.

Finally, the proposed technique RU-BiLSTM achieved the highest accuracy of 0.9245 for positive class and 0.8977 for negative class. Its precision is 0.9537 for positive class and 0.9599 for negative class, while its recall is 0.8837 for positive class and 0.9227 for negative class. Its F1-score is 0.9189 for positive class and 0.9227 for the negative class.

Figures 4.1, 4.2, 4.3 and 4.4 show the experimental results of Long dataset such as Kindle, App, CD and Electronic datasets. It is compared with existing algorithms. Among all these methods, the suggested method achieves greater performance in terms of accuracy, precision, recall, and f1-score.

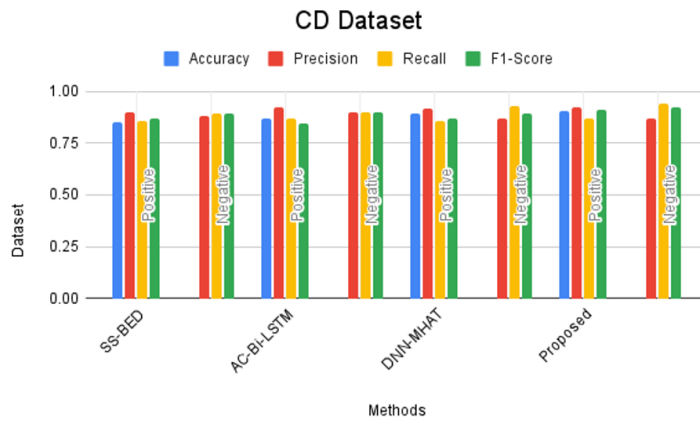


Fig. 4.3: Results of CD dataset

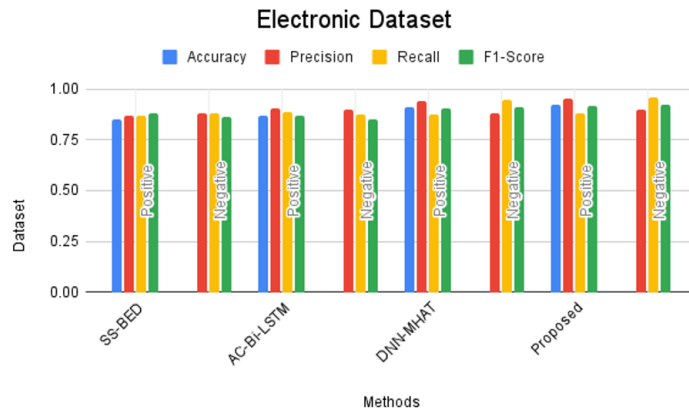


Fig. 4.4: Results of Electronic Dataset

**4.3. Evaluation of Short Tweets.** In short tweets, two types of datasets were used as Airline Twitter dataset and the Sentiment 140 dataset. The positive and negative aspects were classified. Table 4.6 and 4.7 shows the experimental outcome of two tweet datasets.

The table 4.7 displays the experimental results of four different methods applied to the Sentiment 140 dataset, which is a collection of tweets labeled as positive or negative. The methods used are SS-BED, AC-Bi-LSTM, DNN-MHAT, and Proposed RU-BiLSTM, and the evaluation metrics used are Accuracy, Precision, Recall, and F1-Score for both positive and negative classes.

The results show that the Proposed RU-BiLSTM method outperforms the other methods in terms of accuracy, precision, recall, and F1-Score for both positive and negative classes. It achieved an accuracy of 0.8732 for positive and 0.8967 for negative classes, which are the highest among all the methods. The precision, recall, and F1-Score for positive and negative classes are also high, indicating that the Proposed RU-BiLSTM method is effective in classifying tweets as positive or negative.

In Figures 4.5 and 4.6 demonstrate how well RU-NiLSTM performed in terms of accuracy on the Sentiment140 and Airline Twitter datasets, with accuracy scores of 0.39 and 0.45, correspondingly. Upon this Sentiment140 and Airline Twitter datasets, the enhancements for the F1 scale are 0.72% and 0.63% for the positive classes and 0.37% and 0.53% for the negative classes, correspondingly. As shown before, our RU-BiLSTM performed much better than the other techniques in terms of accuracy and F1 scale.

Table 4.6: Experimental outcomes of Airline Twitter Dataset

Techniques Used	Classes Used	Accuracy	Precision	Recall	F1-Score
SS-BED	Positive	0.8789	0.8211	0.8748	0.8673
	Negative		0.8518	0.8312	0.8307
AC-Bi-LSTM	Positive	0.8914	0.8778	0.8714	0.8504
	Negative		0.9072	0.8741	0.8120
DNN-MHAT	Positive	0.9302	0.9588	0.9608	0.9603
	Negative		0.8411	0.8167	0.8261
Proposed RU-BiLSTM	Positive	0.9475	<b>0.9598</b>	<b>0.9789</b>	<b>0.9752</b>
	Negative		<b>0.8567</b>	<b>0.8321</b>	<b>0.8467</b>

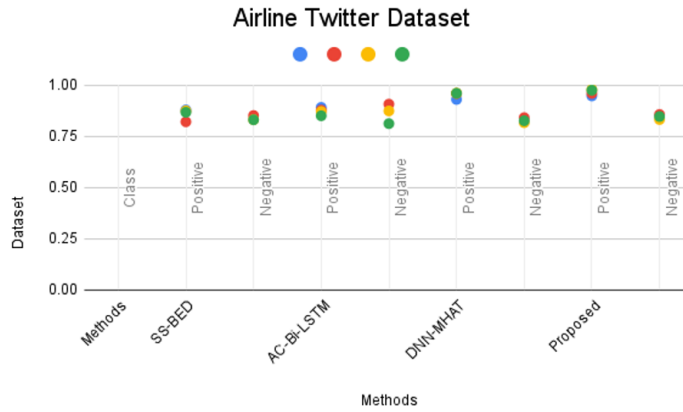


Fig. 4.5: Results of Airline Twitter Dataset

**5. Conclusion.** In this study, we present a deep recurrent Bi-directional Long-short Term Memory model for an aspect-level-based attention mechanism for text sentiment analysis. Initially, the data is collected and then preprocessed using Tokenization, Stop word removal, Stemming, and Lemmatization. After pre-processing, the features are vectorized by using Bag-of-Words and TF-IDF. Then, the extracted features are given into word embeddings by GloVe and Word2vec. It uses Deep Recurrent based Bidirectional Long Short Term Memory (RU-BiLSTM) for aspect-based sentiment analysis. The RU-Bi-LSTM method integrates aspect-based embeddings and an attention mechanism for text classification. Finally, the binary and ternary classification outcomes are obtained using the final dense softmax output layer. In this proposed method, four long review dataset and two short Twitter dataset is used. It classifies the positive and negative aspects based texts. Lastly, it was discovered that the RU-BiLSTM model developed in this work greatly outperformed and compared with DNN-MHAT, AC-BiLSTM, and SS-BED. The proposed RU-BiLSTM outperforms better in terms of accuracy, precision, recall, and f1-score. The proposed model is trained and evaluated only on English text datasets, and it may not be directly applicable to other languages. Further, the model only focuses on aspect-level sentiment analysis and may not be suitable for document-level sentiment analysis. In the future, The proposed model can be extended to support multilingual text classification to improve its applicability to other languages. Also, the proposed model can be extended to incorporate other types of attention mechanisms such as self-attention and transformer-based attention.

## REFERENCES

- [1] M. AL-AYYOUB, A. A. KHAMASEH, Y. JARARWEH, AND M. N. AL-KABI, *A comprehensive survey of arabic sentiment analysis*,



Table 4.7: Experimental results of Sentiment 140 Dataset

Methods Used	Class	Accuracy	Precision	Recall	F1-Score
<b>SS-BED</b>	Positive	0.8789	0.8311	0.8048	0.7773
	Negative		0.7518	0.8312	0.8307
<b>AC-Bi-LSTM</b>	Positive	0.7814	0.7478	0.8714	0.7904
	Negative		0.8172	0.8741	0.8120
<b>DNN-MHAT</b>	Positive	0.8217	0.7782	0.7585	0.8092
	Negative		0.8896	0.7478	0.8212
<b>Proposed RU-BiLSTM</b>	<b>Positive</b>	<b>0.8732</b>	<b>0.8098</b>	<b>0.9137</b>	<b>0.8542</b>
	<b>Negative</b>		<b>0.8967</b>	<b>0.7721</b>	<b>0.8367</b>

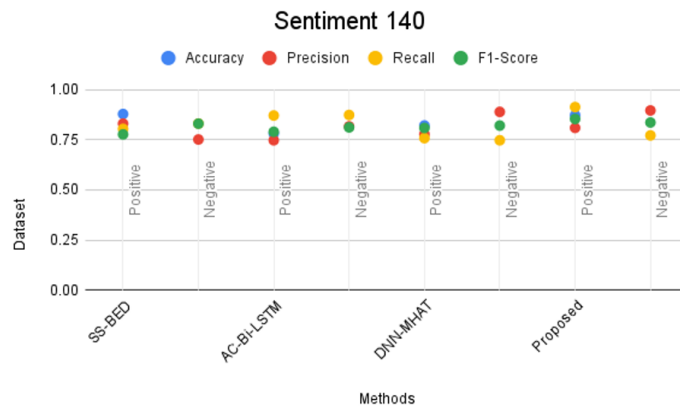


Fig. 4.6: Results of Sentiment140

- Information processing & management, 56 (2019), pp. 320–342.
- [2] M. ALAM, F. ABID, C. GUANGPEI, AND L. YUNRONG, *Social media sentiment analysis through parallel dilated convolutional neural network for smart city applications*, Computer Communications, 154 (2020), pp. 129–137.
- [3] A. H. ALAMOUDI, B. ZAIDAN, M. AL-MASAWA, S. M. TARESH, S. NOMAN, I. Y. AHMARO, S. GARFAN, J. CHEN, M. A. AHMED, A. ZAIDAN, ET AL., *Multi-perspectives systematic review on the applications of sentiment analysis for vaccine hesitancy*, Computers in Biology and Medicine, 139 (2021), p. 104957.
- [4] N. AYOUBI, D. SHARIFRAZI, R. ALIZADEHSANI, A. SHOEIBI, J. M. GORRIZ, H. MOOSAEI, A. KHOSRAVI, S. NAHAVANDI, A. G. CHOFREH, F. A. GONI, ET AL., *Time series forecasting of new cases and new deaths rate for covid-19 using deep learning methods*, Results in physics, 27 (2021), p. 104495.
- [5] P. BALA ANKI REDDY, S. JAKEER, H. THAMEEM BASHA, S. R. REDDISEKHAR REDDY, AND T. MAHESH KUMAR, *Multi-layer artificial neural network modeling of entropy generation on mhd stagnation point flow of cross-nanofluid*, Waves in Random and Complex Media, (2022), pp. 1–28.
- [6] M. BIRJALI, M. KASRI, AND A. BENI-HSSANE, *A comprehensive survey on sentiment analysis: Approaches, challenges and trends*, Knowledge-Based Systems, 226 (2021), p. 107134.
- [7] B. A. CHANDIO, A. S. IMRAN, M. BAKHTYAR, S. M. DAUDPOTA, AND J. BABER, *Attention-based ru-bilstm sentiment analysis model for roman urdu*, Applied Sciences, 12 (2022), p. 3641.
- [8] A. CHATTERJEE, U. GUPTA, M. K. CHINNAKOTLA, R. SRIKANTH, M. GALLEY, AND P. AGRAWAL, *Understanding emotions in text using deep learning and big data*, Computers in Human Behavior, 93 (2019), pp. 309–317.
- [9] Y. CHENG, L. YAO, G. ZHANG, T. TANG, G. XIANG, H. CHEN, Y. FENG, AND Z. CAI, *Text sentiment orientation analysis of multi-channels cnn and bigru based on attention mechanism*, J. Comput. Res. Dev, 57 (2020), pp. 2583–2595.
- [10] A. DE HARO-GARCÍA, G. CERRUELA-GARCÍA, AND N. GARCÍA-PEDRAJAS, *Ensembles of feature selectors for dealing with class-imbalanced datasets: A proposal and comparative study*, Information Sciences, 540 (2020), pp. 89–116.
- [11] Z. GAO, Z. LI, J. LUO, AND X. LI, *Short text aspect-based sentiment analysis based on cnn+ bigru*, Applied Sciences, 12 (2022), p. 2707.
- [12] K. GARCIA AND L. BERTON, *Topic detection and sentiment analysis in twitter content related to covid-19 from brazil and the usa*, Applied soft computing, 101 (2021), p. 107057.

- [13] A. GHAFOR, A. S. IMRAN, S. M. DAUDPOTA, Z. KASTRATI, R. BATRA, M. A. WANI, ET AL., *The impact of translating resource-rich datasets to low-resource languages through multi-lingual text processing*, IEEE Access, 9 (2021), pp. 124478–124490.
- [14] Z. HAMEED AND B. GARCIA-ZAPIRAIN, *Sentiment classification using a single-layered bilstm model*, Ieee Access, 8 (2020), pp. 73992–74001.
- [15] F. HEMMATIAN AND M. K. SOHRABI, *A survey on classification techniques for opinion mining and sentiment analysis*, Artificial intelligence review, 52 (2019), pp. 1495–1545.
- [16] T. HOU AND L. ZHOU, *Chinese ship fault relation extraction method based on bidirectional gru neural network and attention mechanism*, Comput. Sci, 48 (2021), pp. 154–158.
- [17] A. S. IMRAN, S. M. DAUDPOTA, Z. KASTRATI, AND R. BATRA, *Cross-cultural polarity and emotion detection using sentiment analysis and deep learning on covid-19 related tweets*, Ieee Access, 8 (2020), pp. 181074–181090.
- [18] Z. KASTRATI, F. DALIPI, A. S. IMRAN, K. PIREVA NUCI, AND M. A. WANI, *Sentiment analysis of students' feedback with nlp and deep learning: A systematic mapping study*, Applied Sciences, 11 (2021), p. 3986.
- [19] H. KWON AND S. LEE, *Textual backdoor attack for the text classification system*, Security and Communication Networks, 2021 (2021), pp. 1–11.
- [20] A. LIGTHART, C. CATAL, AND B. TEKINERDOGAN, *Systematic reviews in sentiment analysis: a tertiary study*, Artificial Intelligence Review, (2021), pp. 1–57.
- [21] S. MAHESWARAN, S. SATHESH, P. PRIYADHARSHINI, AND B. VIVEK, *Identification of artificially ripened fruits using smart phones*, in 2017 international conference on intelligent computing and control (I2C2), IEEE, 2017, pp. 1–6.
- [22] S. MAHESWARAN, S. SATHESH, G. SARAN, AND B. VIVEK, *Automated coconut tree climber*, in 2017 International Conference on Intelligent Computing and Control (I2C2), IEEE, 2017, pp. 1–6.
- [23] K. MEHMOOD, D. ESSAM, K. SHAFI, AND M. K. MALIK, *Discriminative feature spamming technique for roman urdu sentiment analysis*, IEEE Access, 7 (2019), pp. 47991–48002.
- [24] A.-S. MOHAMMAD, M. M. HAMMAD, A. SA'AD, A.-T. SAJA, AND E. CAMBRIA, *Gated recurrent unit with multilingual universal sentence encoder for arabic aspect-based sentiment analysis*, Knowledge-Based Systems, (2021), p. 107540.
- [25] A. B. NASSIF, A. ELNAGAR, I. SHAHIN, AND S. HENNO, *Deep learning for arabic subjective sentiment analysis: Challenges and research opportunities*, Applied Soft Computing, 98 (2021), p. 106836.
- [26] S. RANI AND P. KUMAR, *Deep learning based sentiment analysis using convolution neural network*, Arabian Journal for Science and Engineering, 44 (2019), pp. 3305–3314.
- [27] Z. SAFDAR, R. S. BAJWA, S. HUSSAIN, H. B. ABDULLAH, K. SAFDAR, AND U. DRAZ, *The role of roman urdu in multilingual information retrieval: A regional study*, The Journal of Academic Librarianship, 46 (2020), p. 102258.
- [28] H. S. SHARAF AL-DEEN, Z. ZENG, R. AL-SABRI, AND A. HEKMAT, *An improved model for analyzing textual sentiment based on a deep neural network using multi-head attention mechanism*, Applied System Innovation, 4 (2021), p. 85.
- [29] B. VIVEK, S. MAHESWARAN, P. KEERTHANA, S. SATHESH, S. BRINGERAJ, R. A. SRI, AND S. A. SULTHANA, *Low-cost raspberry pi oscilloscope*, in 2018 international conference on intelligent computing and communication for smart World (I2C2SW), IEEE, 2018, pp. 386–390.
- [30] L. WANG, C. LIU, D. CAI, AND T. LU, *Chinese text sentiment analysis based on cnn-bigru network with attention mechanism*, Journal of Computer Applications, 39 (2019), p. 2841.
- [31] W. WANG, Y. SUN, Q. QI, AND X. MENG, *Text sentiment classification model based on bigru-attention neural network*, Appl. Res. Comput, 36 (2019), pp. 3558–3564.
- [32] A. YADAV AND D. K. VISHWAKARMA, *Sentiment analysis using deep learning architectures: a review*, Artificial Intelligence Review, 53 (2020), pp. 4335–4385.
- [33] S. YAO, P. JING, P. LI, AND J. CHEN, *A multi-dimension traceable privacy-preserving prevention and control scheme of the covid-19 epidemic based on blockchain*, Connection Science, 34 (2022), pp. 1654–1677.
- [34] L. YUE, W. CHEN, X. LI, W. ZUO, AND M. YIN, *A survey of sentiment analysis in social media*, Knowledge and Information Systems, 60 (2019), pp. 617–663.
- [35] B. ZHANG, X. ZHOU, Y. LUO, H. ZHANG, H. YANG, J. MA, AND L. MA, *Opportunities and challenges: Classification of skin disease based on deep learning*, Chinese Journal of Mechanical Engineering, 34 (2021), pp. 1–14.
- [36] X. ZHANG, M. ZHANG, X. LI, ET AL., *Rolling bearing fault mode recognition based on 2d image and cnn-bigru*, J. Vib. Shock, 40 (2021), pp. 194–201.

*Edited by:* Sathishkumar V E

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

*Received:* Feb 26, 2023

*Accepted:* May 2, 2023



## A DEEP LSTM-RNN CLASSIFICATION METHOD FOR COVID-19 TWITTER REVIEW BASED ON SENTIMENT ANALYSIS

**Abstract.** In today's world, advanced internet technologies have significantly increased people's affinity towards social networks to stay updated on current events and communicate with others residing in different cities. Social opinion analyses helped determine the optimal public health response during the COVID-19 pandemic. Analysis of articulating tweets from Twitter can reveal the public's perceptions of social distance. Sentiment Analysis is used for classifying text data and analyzing people's emotions. The proposed work uses LSTM-RNN with the SMOTE method for categorizing Twitter data. The suggested approach uses increased characteristics weighted by attention layers and an LSTM-RNN-based network as its foundation. This method computes the advantage of an improved information transformation framework through the attention mechanism compared to existing Bi-LSTM and LSTM models. A combination of four publicly accessible class labels such as happy, sad, neutral, and angry, is analyzed. The message of tweets is analyzed for polarization and subjectivity using TextBlob, VADER (Valence Aware Dictionary for Sentiment Reasoning), and SentiWordNet. The model has been successfully built and evaluated using two feature extraction methods, TF-IDF (Term Frequency-Inverse Document Frequency) and Bag of Words (BoW). Compared to the previous methodologies, the suggested deep learning model improved considerably in performance measures, including accuracy, precision, and recall. This demonstrates how effective and practical the recommended deep learning strategy is and how simple it is to employ for sentiment categorization of COVID-19 reviews. The proposed method achieves 97% accuracy in classifying the text whereas, among existing Bi-LSTM, achieves 88% maximum in the text classification.

**Key words:** Sentiment Analysis, Covid-19, deep learning, Twitter reviews, social networks, classification, SentiWordNet, TextBlob, Bag of Words.

**1. Introduction.** The worldwide pandemic of coronavirus disease (COVID-19) has adverse effects on human health [29]. From its first detection in Wuhan, China, it has already spread to multiple nations on all continents, and on March 11, 2020, the World Health Organization (WHO) declared it a pandemic. On numerous social media platforms, there has been a lot of COVID-19 information shared. Moreover, misleading news might propagate on social media platforms like Twitter; thus, it's crucial to comprehend people's emotions from such textual resources [26]. The identification of attitudes from Twitter data on material about COVID-19 can be made with the help of deep learning algorithms. Moreover, given the difficulties involved in textual analysis, there still needs to be a technological issue regarding how deep learning networks may be adapted and adjusted to obtain high accuracy.

At the beginning of the COVID-19 outbreak, social media quickly became a vital communication method for creating, disseminating, and consuming data. The discovery and characterization of infectious disease outbreaks and understanding people's sentiments, behaviors, and views have all benefited from including social media data in numerous research. Social media user-generated material can be opinionated or inaccurate, frequently including false data and conspiracy theories [17]. The practice of categorizing feelings in qualitative information utilizing machine learning (ML) and natural language processing is known as sentiment analysis, commonly referred to as opinion mining (NLP) [4].

Different methods for evaluating consumer emotion in social media data have been presented in previous research [27, 31, 20, 1, 25]. Propose a hybrid machine learning technique, for instance, to categorize user emotions as either positive or negative. In a collection of customer reviews, the [12] authors employ natural language processing (NLP) to decipher user sentiment. A Bayesian graphical method is used to assess data from Twitter. Such practices may be computationally challenging, time-consuming, and not always offer high precision.

This work provides a deep learning method for categorizing the emotion of tweets about COVID-19. Positive and negative feelings can be distinguished through an analysis of people's opinions, and efforts are being made to detect texts in the literature expressing negative emotions [11, 7]. The rectified linear activation function (ReLU)

is frequently employed because of the low complexity of training and the potential for higher results. Most critically, nonlinear connections are commonly studied using ReLU. Because ReLU can help convolution neural networks better capture complicated patterns, it has been used in several research neural network models [8].

Several frameworks, such as the Long-Short Term Model (LSTM), have been suggested to sidestep the constraints of neural networks. LSTM has been recognized as a crucial component because it effectively solves time-series data and sequentially issues [19]. Due to its capacity to understand text sequence and identify relationships among words or phrases in sentiment analysis, the suggested deep learning strategy incorporates the LSTM non-linear activation [14]. The main objective of the research is to classify the sentiment of the text with high accuracy. The research question relies on

1. What is the effectiveness of using LSTM-RNN with SMOTE method and attention layers for sentiment analysis of COVID-19-related tweets, compared to previous methodologies using TextBlob, VADER, and SentiWordNet?
2. How the sentimental analysis is performed effectively on the covid 19 twitter data?

The proposed method uses LSTM- recurrent neural network-based SMOTE for classifying the Covid-19 Twitter reviews.

The main contribution of the proposed method is given below:

1. Utilizing attentive learning and LSTM, the primary aim is to maximize effective weight by the semantic relevance of terms.
2. To evaluate the effectiveness of TF-IDF, BoW, and feature extraction methods, as well as the results of various deep learning models for sentiment analysis using different annotation techniques like TextBlob, VADER, and SentiWordNet.
3. SMOTE helps to reduce the imbalanced issues of datasets. also, with the help of the majority class, it balances the datasets.

The rest of our research article is written as follows: Section 2 discusses the related work on COVID-19 Twitter reviews and Deep learning classification methods. Section 3 shows the general working methodology of the proposed work. Section 4 evaluates the implementation and results of the proposed method. Section 5 concludes the work and discusses the result evaluation.

**2. Related Works.** The COVID-19 pandemic has caused massive deaths of innocent people throughout the globe and constitutes a severe risk to the food supply, the workplace, and public health. Numerous people have died due to the COVID-19 epidemic, which also poses a severe threat to workplace safety, food production, and the promotion of healthy lifestyles. The characteristics of COVID-19 comprise death, transmission characteristics, period of national viruses, and initial fatalities. This is exposed by the disparities in social networking and global financial responses due to the extreme viral spread's residual effects. As users utilized the channels to vent their feelings during the lockdown, social networking sites were crucial in sharing data well about epidemics around the globe. Given this critical scenario, it is essential to look at how people are responding on Twitter while considering standard terms relevant to the outbreak [5].

To read or understand all that's been stated on Twitter about COVID-19 immunizations would only be possible for a person. Moreover, we can utilize natural language processing (NLP) techniques to examine an extraordinarily complex and wide-ranging discourse utilizing word cloud visualizations, sentiment analysis, and linguistic extraction of features. Using deep learning classifiers, the author conducted a study on the sentiment classification of COVID-19 tweets [6]. This study found that tweets regarding the COVID-19 pandemic did not influence people. The results demonstrate that neither the number of words in tweets nor WordCloud includes helpful comments. The assertions are validated by a suggested deep learning classifier model, which has an accuracy of up to 81 percent. According to the researchers, a fuzzy approach built on Gaussian participation may accurately identify tweet attitudes.

Using the BERT model, the author has suggested conducting a sentiment analysis of how the coronavirus affects social interactions. They assert that since Twitter has grown to be one of the most well-known social media platforms, the authors used the BERT model to conduct sentiment analysis on Twitter to comprehend people's emotions and psychological states better [28]. In this study, the authors performed a sentiment classification on two variables: the first group had Twitter from users around the globe, and the other set contained tweets from users in India. They checked the reliability of the emotion classification using the GitHub repository.

Table 2.1: Previous studies based on Covid-19 Twitter reviews

Author	Methodology Used	Dataset	Objective
Relucio, F.S.; Palaoag, T.D [22]	Web analytics strategy	There are 1717 tweets, all were gathered on Twitter	Observe the attitude of informative posts.
Bhagat, K.K.; Mishra, S. [3]	Using a thesaurus as a guide	There are 154 documents gathered from Google	Public perception of COVID-19's online education.
Ashwitha, R.; Jeevitha [1]	NLP methods and a classification using logistic regression	Data collection is done using Google Forms.	Evaluation of Public Opinion on the COVID-19 Epidemic's Teaching
Althagafi, A.; Althobaiti, [2]	Random forest, KNN, and Naive Bayes	The Twitter API was used to collect 10,445 Twitter posts.	Coronavirus sentiment classification in online learning
Mamtesh, M.; Mehla, [16]	Logistic Regression and KNN	Evaluations of movies with emotional commentary	The information was gathered from a multitude of sources.
Imran, A.S.; Daudpota, S.M. [24]	In-depth learning (Multi-layer LSTM)	There are 27,357 tweets in total relating to COVID-19	COVID-19 sentiment classification
Chintalapudi, N.; Battineni, G. [10]	In-depth learning (BERT and LSTM)	There were 3090 tweets about COVID-19 overall.	COVID-19 sentiment classification

The experimental results show that the accuracy rate is 94%.

**3. Proposed Twitter Analysis Methodology .** The proposed method uses LSTM-RNN with SMOTE for classifying the twitter text data. Initially, it collects the data from the Twitter reviews and then the collected data is preprocessed by using tokenization, URL removal, stop word removal, stemming, and lemmatization. Next, the TextBlob is used to compute the polarity score, and then the feature engineering process is done. Next, the LSTM-RNN with SMOTE method is used for assigning efficient weights. Figure 3.1 shows the architecture of the proposed method. first the twitter data is collected and preprocessed by removing URL, Noises, spaces, etc from the dataset. Text Blob is based on the Naive Bayes algorithm and is suitable for analyzing short and informal texts, such as social media posts and product reviews. This helps to classify text polarity of the context. further the features are extracted and analyzed using frequency of the terms and bag of words. Feature matrix is computed by the training and testing the features. Finally, classification is done using LSTM-RNN with SMOTE to obtain the optimal results.

**3.1. Data Collection.** On Twitter, companies may engage with customers in much more intimate ways. But choosing which tweets to be responded to first can be challenging for marketers because Twitter contains so much information. Sentiment analysis has become necessary as a technique in social media management initiatives. Using sentiment analysis technology, social media activities are automatically monitored for emotion. The initial stage in solving the topic of sentiment classification of a Covid 19 tweets is information gathering [18]. We took advantage of Tweepy and Twitter's streaming API to follow specific popular phrases and profiles during information gathering. Utilizing the Twitter API to create the large dataset entails a sequence of iterations. This involves setting up an identity, installing Tweepy, running a small test, looking at a tweet's JSON file, extracting the information, and gathering the data. Following data collection, preprocessing must be done before moving on.

**3.2. Data Pre-processing.** Applications that analyze the data must pre-process the data to eliminate extraneous information and speed up the learning of categorization algorithms. Any material that makes almost no contributions to determining the desired class is referred to as redundant data, it expands the feature

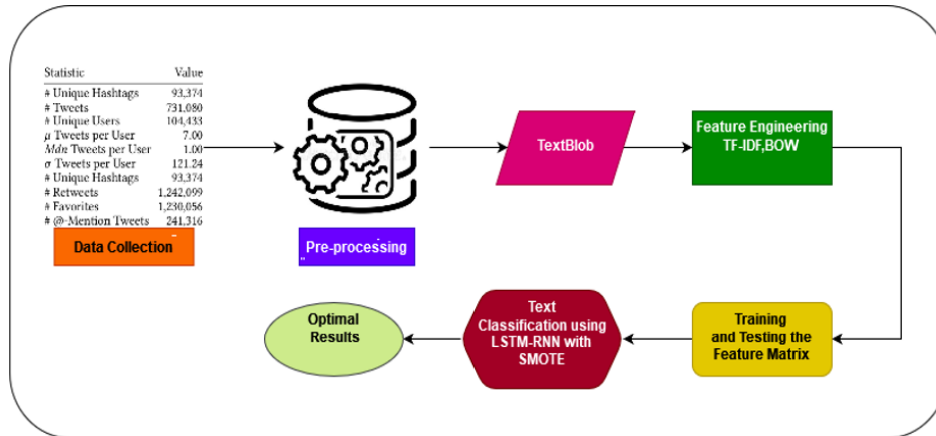


Fig. 3.1: Architecture of Proposed Method

representation and adds needless computing complexity. As a result, classification techniques perform worse when preprocessing is neglected or done incorrectly. So, before preprocessing, data screening or pretreatment is done.

The second phase includes both exploratory data analysis and data pre-processing. Once the data had been cleaned to extract meaningful information, the raw tweets could not produce objective findings for Sentiment predictions. The major obstacles were #tags, @mentions, URLs, and stop words in tweets. The #tags, @mentions, and URLs again from tweets are substituted using regular expressions. Stop words are eliminated using the Python NLTK library.

**3.2.1. URL removals.** Although they have yet to add new definitions for training systems to simplify the feature space, URL links, tags, punctuation, and numerals do not improve classifier performance. By eliminating them, the feature representation is made simpler.

**3.2.2. Stop words removal.** Stop words are often used yet provide no meaningful data for the research. Stop words like "the," "is," "a," and "an" are eliminated.

**3.2.3. Stemming and Lemmatization.** While stemming and lemmatization focus on reducing a word's inflectional forms and, occasionally, its derivationally related forms to a basic shape. In this way, words like "walks," "walking," and "walked" are changed to the core word "walk."

**3.3. TextBlob.** The lexicon-based TextBlob method can be applied to various Natural language processing tasks, such as part-of-speech labeling, sentiment classification, noun word retrieval, paraphrasing, and grouping [30]. It was employed in this investigation for sentimental reasons. A polarization score between one and one is provided by the TextBlob emotion method. Less than zero indicates a negative comment, zero indicates a reasonable statement, and more than zero shows a positive assertion in a tweet [23]. VADER uses a rule-based approach to analyze the polarity of text and can handle the complexity of social media language, such as slang, emoticons, and abbreviations. VADER provides not only the text's polarity but also the sentiment's intensity, making it useful for fine-grained sentiment analysis. SentiWordNet uses a hybrid approach that combines machine learning and lexicon-based methods to analyze the sentiment of text. SentiWordNet helps analyze longer texts and can provide a more detailed analysis of the sentiment by considering the sentiment of each word in context.

**3.4. Feature Engineering.** The second most prominent feature extraction techniques, BoW and TF-IDF, are employed to collect features from tweets.

**3.4.1. Bag of Words.** The boW is a straightforward method for extracting characteristics from condensed texts or information and is frequently used in information retrieval and natural language processing. The boW

should be used to estimate the appearance of a phrase in a word and create a feature representation that includes the frequency of each particular comment. The primary purposes of the BoW are to expand the dictionary of all unique terms and to train the teaching methods using their frequency distributions.

**3.4.2. Term Frequency-Inverse Document Frequency.** For extracting the features, a TF-IDF method is employed to collect weighted features using textual information. To help learning models perform much better, it offers the value of every phrase in the library. The result of TF and IDF is TF-IDF. TF is calculated by using the following:

$$TF(t, d) = \frac{n_t}{N_{T,d}} \quad (3.1)$$

Here  $N_{T,d}$  denotes the set of terms  $T$  in the documents, and  $n_t$  is the number of times the term  $t$  appears in content  $d$ . The IDF of a word reveals its importance over the entire set and could be determined as follows:

$$IDF = \log \frac{Doc}{n_d} \quad (3.2)$$

Here  $n_d$  is the range of resources where its term  $t$  occurs, and  $Doc$  is the maximum number of documents in the collection. The formula for calculating TF-IDF utilizing TF and IDF is

$$TF-IDF = TF \cdot IDF \quad (3.3)$$

### 3.5. Text Classification using LSTM-RNN with SMOTE method.

**3.5.1. SMOTE.** By rebalancing the number of observations for each class in a database, SMOTE (Synthetic Minority Oversampling Technique) is utilized to overcome the challenges associated with imbalanced datasets. Through creating synthetic examples of minority classes, it is possible to establish balance by matching the quantity of minority class samples with those of the majority class label.

Following using TextBlob, the proportion of feelings is not identical, which makes it possible for algorithms to generalize the skewed information. SMOTE is used to equalize the dataset by producing inaccurate numbers for the minority class to prevent this over-fitting issue.

**3.5.2. Training and Testing the text classification using LSTM-RNN with SMOTE.** This study aims to investigate sentiment classification accuracy using tweet data using a deep learning architecture with attention levels. There are many additional processing stages for the LSTM-RNN with SMOTE architecture that is presented. The convolution of features comes next. The LSTM-RNN model is fed the input COVID-19 tweets throughout this stage. This phase aims to identify significant lexical elements from the sentence construction. Additionally, the LSTM-RNN method generates feature vectors and determines the temporal link between features. To minimize the features overlap, the weights provided by the LSTM-RNN mapping characteristics are modified after the feature mapping has been completed. Utilizing the learning algorithm to improve weights aids in the selection of pertinent details. Figure 3.2 shows the workflow of the proposed method.

To train the dataset, we also take into account the semantic information of the tweets and produce a second set of keywords with numbers attributed to the emotions conveyed in those tweets: afraid = 0, sadness = 1, angry = 2, and happiness = 3. Recognizing that this allocation of emotion labels is arbitrary, the goal is to produce a matrices level that can be input to LSTM-RNN to produce results for sentiment classification.

The RNN model recognizes the series of pixels  $P_x = PX_1, PX_2, \dots, PX_n$ , generates hidden layer  $HI = HI_1, HI_2, \dots, HI_n$ , and generates output states  $OT = OT_1, OT_2, \dots, OT_n$  in the manner shown in Figure 3.3.

$$OT_t = \sigma(WE_{HI_t OT_t} + b_t) \quad (3.4)$$

$$HI_t = \sigma(WE_{HI_{t-1} HI_t} HI_{t-1} + WE_{PX_t HI_t PX_t} + b_{HI_t}) \quad (3.5)$$

In this case,  $WE_{HI_t OT_t}$  stands for the vector from the hidden unit  $HI_t$  and the output unit  $OT_t$ ,  $HI_{t1}$  for the hidden unit for a  $t-1$  pixel series,  $WE_{HI_t HI_t}$  for the sequencing period  $t$ , and  $b_{HI_t}$  and  $b_t$  for bias.

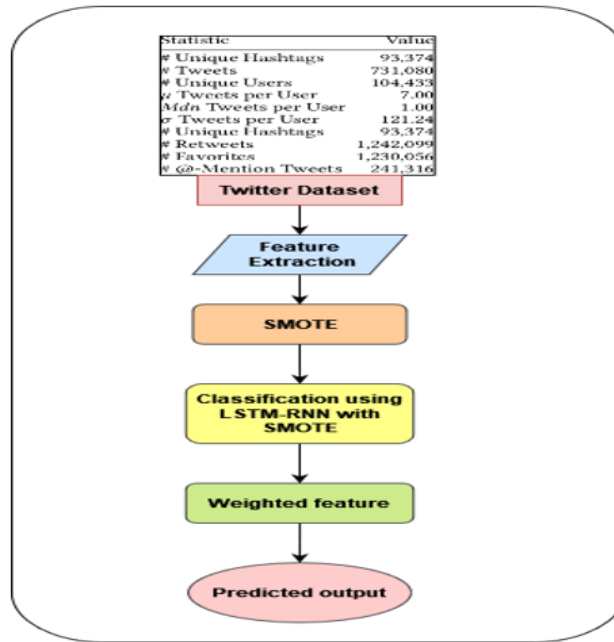


Fig. 3.2: Workflow of Proposed Method

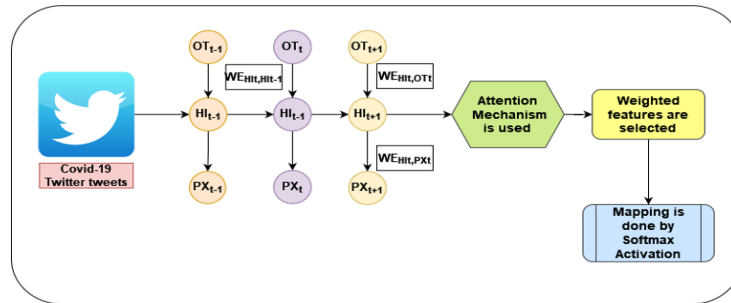


Fig. 3.3: Structure of RNN

The LSTM-enhanced RNN's portion is shown in a modeling framework for RNN in Figure 3.3 above, along with suggestions made. Additionally, the LSTM stacking will be used to train the timing series characteristics in issues containing a single range of views, which is necessary for a model to learn from in ability to forecast the subsequent value in a sequence.

$$ig_t = \tanh(WE_{PX_t ig_t} PX_t + WE_{HI_{t-1} ig_t} HI_{t-1} + b_{ig_t}) \quad (3.6)$$

$$PX_t = \sigma(WE_{PX_t PE_t} PX_t + WE_{HI_{t-1} PE_t} HI_{t-1} + b_{pX_t}) \quad (3.7)$$

$$fg_t = \sigma(WE_{PX_t fg_t} PX_t + WE_{HI_{t-1} fg_t} HI_{t-1} + b_{fg_t}) \quad (3.8)$$

$$OTPX_t = \sigma(WE_{PX_t OTPX_t} PX_t + WE_{HI_{t-1} OTPX_t} HI_{t-1} + b_{OTPX_t}) \quad (3.9)$$



$$CE_t = ce_{t-1} \Theta f_{g_t} + ig_t \Theta PX_t \quad (3.10)$$

$$HI_t = \tanh(CE_t) \Theta OTPX_t \quad (3.11)$$

Here  $ig_t$  stands for the input nodes,  $PX_t$  for predictions in the first levels,  $f_{g_t}$  for forget gate,  $H_t$  for output data,  $b_{ig_t}$ ,  $b_{PX_t}$ ,  $b_{f_{g_t}}$ , and  $b_{OTPX_t}$  for bias vectors,  $Ce_t$  for cell state, and  $WE_{PX_t f_{g_t}}$  for weight matrices. To retrieve linguistic information of the input tweets, it is possible to mix both RNN and LSTM systems.

Importantly, as mentioned in [15], we use attention layer in the research to enhance learning from features and feature weights. The LSTM-RNN is being used to generate features that are weighted by the attentiveness process to understand phrase sequences. Additionally, the usage of secondary labeling in conjunction with LSTM-RNN makes it easier for learners to gain more technical knowledge. The concentration functional of the attention layer analyzes distribution of weight and calculates arrays for the various levels, according to the research [21].

As a result, given an input  $PX_i$ , the characteristics  $f(PX_i, PX_{i+1})$  are those produced from the second layer, and  $f(PX_i, PX_{i+1}, \dots, PX_i + L-1)$  are those produced from the  $L$ -th layer. These features values represent the reactions of multi-scale  $n$ -grams namely the unigram  $PX_i$ , bigram  $PX_i PX_{i+1}$ , and  $L$ -gram  $f(PX_i PX_{i+1} \dots PX_i + L-1)$ . The filtering ensembles and reweight scale were utilized together in the focusing process. Additionally, taking the descriptors as input, scale reweight is employed to calculate the SoftMax distribution of attention weights and produces weighed feature values for suggests that the market [9, 13].

$$S_i^i = FEL_{ensm}(PX_i^i) \quad (3.12)$$

$$PX_{atten}^i = \sum_{j=1}^L \alpha_L^i PX_L^i \quad (3.13)$$

$$\partial_i^L = Softmax(MLP(PX_{atten}^i)) \quad (3.14)$$

Accuracy, precision, and recall are the three performance criteria used to evaluate the effectiveness of the suggested deep learning model. The originality of the suggested strategy lies in the enhancement of feature weighting achieved through the attention layering method.

The suggested method extracts textual information from an LSTM-RNN series that has been SMOTE-mapped; the LSTM creates a series of comments for every inputs. The vectors employed in this study are essentially a concatenated of the encoder's hidden layers, as well as the attention layer technique then refines the characteristics. The softmax activation mechanism significantly enhances the feature representation assistance provided by the attention mechanism.

**4. Result Analysis.** The Python language and the TextBlob and Tweepy libraries are used in the study to examine the emotion of Twitter posts. Employing phrases, tags, tweets, patterns, or geolocation, Tweepy enables us to locate pertinent information. The following command is used to search tweets that have the COVID hash tag. For instance the twitter posts are given below [29].

```
api = tweepy.API(auth, wait_on_rate_limit=True) #important
```

```
self.tweets = tweepy.Cursor(api.search, q='COVID19, lang="en").items \quad (4.1)
```

The trials made use of the prepared dataset, which had classes that were positive, negative, and neutral. 50,000 tweets in all, 15,000 of which were used for the testing cycle, were used during the training phase. The suggested method combines a number of classifiers such as LSTM-RNN, Bi-LSTM and CNN-LSTM. Table 2.1 shows the total number of Covid-19 tweets used for training and testing.

We conducted several experiments to find the parameters for the pre-processing methods and classifiers that would produce most precise findings. The selected pre-processing techniques are fully detailed in Section

Table 4.1: Description of the Covid19 tweets dataset's

Characteristics Used	Total number of Tweets	Unique Tweets	Tweets (%)
Hashtag #	4317928	683108	40%
Mention @	6942379	1351963	50%
Entities	12826437	385273	80%

Table 4.2: Instances based on Covid-19 tweets

Instances based on Covid-19 tweets	Emotional Kind
Medical centre for Good Visions is moving all of its patients to make a way for healthy COVID-19 situations.	Mixed opinions
Anyone who supports Boris getting the Corona Virus is a total cunt, even a fellow nationalist.	Negative or Sad
Hello, Twittizens. I hope your day is free of coronas.	Positive or Happy
The coronavirus might be prevented from entering if I shut my front doors.	Angry
Attackers make fake coronavirus maps to spread malware to visitors.	Afraid
The thought of Jacob's Nashville show getting postponed makes my heart ache so deeply. Leave now, please.	Negative / Sad

Table 4.3: Experimental Results using Various Classifiers

Classifiers Used	Accuracy (%)	Precision (%)	Recall (%)
Bi-LSTM	88.65%	83%	81%
LSTM-RNN	81.5%	92%	79%
CNN-LSTM	78%	80%	75%
<b>Proposed Method</b>	<b>97.5 %</b>	<b>98%</b>	<b>89%</b>

Table 4.4: Comparative Analysis of Various Activation Function

Activation Function	Accuracy	Precision	Recall
Sigmoid	84.75%	83.3%	82.65%
RELU	85.36%	82.75%	84.25%
TANH	83.69%	80.95%	80
LEAKY RELU	87.32%	83.57%	85.63%

3. In Table 4.2 the sample instances based on Covid-19 tweets and the emotional category. Like these instances were used for training and testing the datasets.

The proposed LSTM-RNN-SMOTE method uses following metrics for evaluating the Covid-19 twitter tweets. The metrics are accuracy, precision, recall, ROC curve and Pie chart for displaying the emotions of people.

Table 4.3 shows the evaluation results of various deep learning classifiers such as LSTM-RNN, Bi-LSTM, CNN-LSTM and Proposed methods are given. Among all the experimental results the proposed method outperforms better.

In figure 4.1 shows the accuracy of various deep learning classifiers. It uses Covid-19 tweets for classifying the emotions. The proposed method LSTM-RNN-SMOTE achieves 97.5% of accuracy in classifying the emotions of people comparing with other classifiers. The Bi-LSTM achieves 88.65%, LSTM-RNN achieves 81.5%

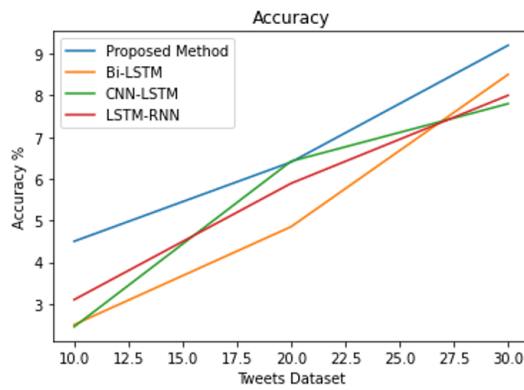


Fig. 4.1: Accuracy

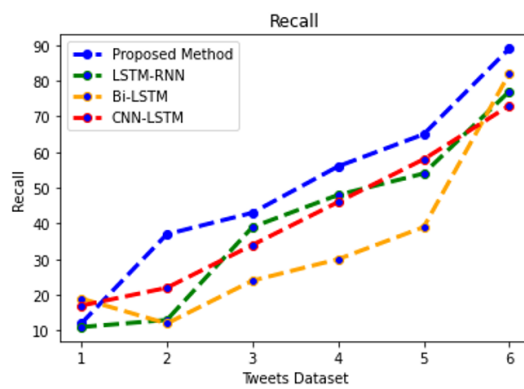


Fig. 4.2: Recall

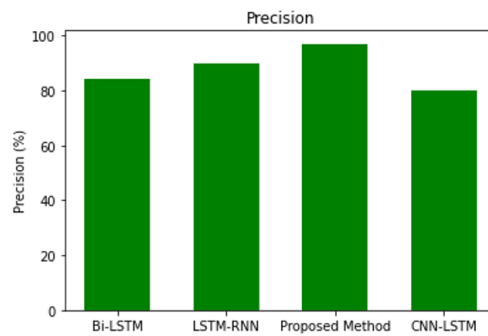


Fig. 4.3: Precision

and CNN-LSTM achieves 78%.

In figure 4.2 shows the recall of various deep learning classifiers. It uses Covid-19 tweets for classifying the emotions. The proposed method LSTM-RNN-SMOTE achieves 89% of precision in classifying the emotions of people comparing with other classifiers. The Bi-LSTM achieves 81%, LSTM-RNN achieves 79% and CNN-LSTM achieves 75%.

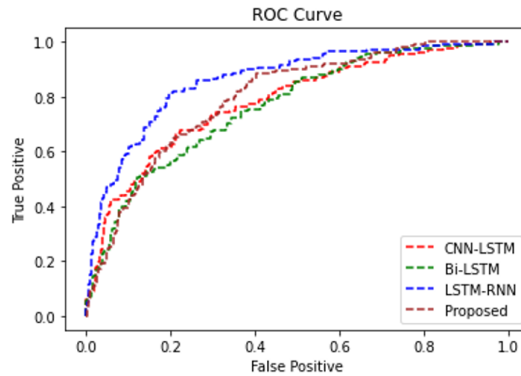


Fig. 4.4: ROC Curve

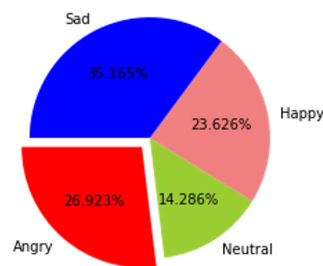


Fig. 4.5: The emotions are classified and displayed in Pie Chart

In figure 4.3 shows the precision of various deep learning classifiers. It uses Covid-19 tweets for classifying the emotions. The proposed method LSTM-RNN-SMOTE achieves 98% of precision in classifying the emotions of people comparing with other classifiers. The Bi-LSTM achieves 83%, LSTM-RNN achieves 92% and CNN-LSTM achieves 80%.

Considering all of these difficulties, the classifiers demonstrated excellent results in differentiating COVID-19 tweets, indicating that current deep learning techniques can generate very efficient and convenient classifiers when used with a Covid-19 tweet dataset. Traditional classifiers do better with greater ROC values than machine learning does. The empirical findings clearly show that classification algorithm can be a useful method for enhancing the efficiency of conventional classifications. In figure 4.4 shows the ROC curve performance.

In figure 4.5 shows the pie chart with the percentage of classified emotions such as Happy, Sad, Neutral and Angry by using Covid-19 tweets. It shows that Happy emotions are 3.626%, Sad emotions are 35.165%, Neutral emotions are 14.286% and Angry emotions are 26.923%. From the results we understand during pandemic situation the people mostly tweets sad emotions and then the angry emotion percentage is higher.

Also, the proposed method uses activation function in attention layer. The activation function findings are shown in Table 4.4, which unmistakably demonstrates that four attention layers with activation function Leaky ReLU offered greater accuracy, precision and recall in compared to certain other activation functions. The accuracy, precision and recall for the Leaky ReLU activation function are shown to be 87.32 percent, 83.57 percent and 85.63 percent respectively. This is significantly greater than the accuracy displayed by ReLU, which is 85.36 percent with a precision and recall of 82.75 percent and 84.25 percent.

As can be shown, Leaky ReLU outperforms other activation functions taken into account TANH, sigmoid, and ReLU in terms of accuracy, precision and recall values. Leaky ReLU is a backpropagation-based activation function that affects both forward and backward learning in LSTM-RNN. The interpretation of the results would depend on the proportion and distribution of each sentiment category in the analyzed text. if the

analysis finds a high proportion of negative sentiment, it may suggest that the public's perceptions of social distancing measures are largely negative and that there is a need for improved communication and public engagement around the issue. Similarly, if there is a high proportion of positive sentiment, it may suggest that the public is generally supportive of social distancing measures and that there may be opportunities to build on this positive sentiment to promote greater compliance with public health guidelines.

**5. Conclusion.** Social networking websites have risen in popularity and are now a powerful way to influence people and educate the common person. Therefore, sentiment analysis for short communications like Twitter is especially difficult because the texts lack relevant information. As an outcome, many methods are constantly being developed to produce the best sentiment model analysis outcome. To finish the classification procedure, a preparatory stage of text pre-processing and extraction of features is necessary. We do numerous tests on various produced databases because pre-processing actions have an effect on classification performance. As a result, this research provides a deep learning method for analyzing the emotions in Twitter messages on COVID-19 opinions. The technique uses enhanced featured weighting from an attention layer and is dependent on an LSTM-RNN-SMOTE network. Through the attention mechanism, this algorithm takes advantage of an improved feature transformation framework. In this investigation, four class labels from a publicly accessible Twitter database (sad, happy, neutral, and angry) were employed. The suggested deep learning strategy improved significantly the performance metrics when compared to existing methods, with gains of 25% in accuracy, 11%–14% in precision, but only 15% in recall. Therefore, the suggested deep learning method for classifying the emotion of COVID-19 evaluations is proven to be effective and practical. The suggested approach is also compared with LSTM-RNN, CNN-LSTM, and Bi-LSTM. The proposed method classifies emotions more efficiently.

In the future, the for-sentiment analysis can be applied to other health-related applications, such as analyzing the sentiment of patient reviews of healthcare providers or identifying negative sentiment around certain health-related products or services.

#### REFERENCES

- [1] M. AFLAKPARAST, M. DE GUNST, AND W. VAN WIERINGEN, *Analysis of twitter data with the bayesian fused graphical lasso*, PloS one, 15 (2020), p. e0235596.
- [2] A. ALTHAGAFI, G. ALTHOBAITI, H. ALHAKAMI, AND T. ALSUBAIT, *Arabic tweets sentiment analysis about online learning during covid-19 in saudi arabia*, International Journal of Advanced Computer Science and Applications, 12 (2021).
- [3] K. K. BHAGAT, S. MISHRA, A. DIXIT, AND C.-Y. CHANG, *Public opinions about online learning during covid-19: a sentiment analysis approach*, Sustainability, 13 (2021), p. 3346.
- [4] M. BHAT, M. QADRI, M. KUNDROO, N. AHANGER, B. AGARWAL, ET AL., *Sentiment analysis of social media response on the covid19 outbreak*, Brain, behavior, and immunity, 87 (2020), p. 136.
- [5] A. K. CHAKRABORTY, S. DAS, AND A. K. KOLYA, *Sentiment analysis of covid-19 tweets using evolutionary classification-based lstm model*, in Proceedings of Research and Applications in Artificial Intelligence: RAAI 2020, Springer, 2021, pp. 75–86.
- [6] K. CHAKRABORTY, S. BHATIA, S. BHATTACHARYYA, J. PLATOS, R. BAG, AND A. E. HASSANIEN, *Sentiment analysis of covid-19 tweets by deep learning classifiers—a study to show how popularity is affecting accuracy in social media*, Applied Soft Computing, 97 (2020), p. 106754.
- [7] S. CHETANPAL, W. SANTOSO, AND G. SRIMANNARAYANA, *A deep learning approach for human face sentiment classification*, in 2021 21st ACIS International Winter Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD-Winter), IEEE, 2021, pp. 28–32.
- [8] J. HAN, K. QIAN, M. SONG, Z. YANG, Z. REN, S. LIU, J. LIU, H. ZHENG, W. JI, T. KOIKE, ET AL., *An early study on intelligent analysis of speech under covid-19: Severity, sleep quality, fatigue, and anxiety*, arXiv preprint arXiv:2005.00096, (2020).
- [9] A. HERNÁNDEZ AND J. M. AMIGÓ, *Attention mechanisms and their applications to complex systems*, Entropy, 23 (2021), p. 283.
- [10] R. S. JAGDALE, V. S. SHIRSAT, AND S. N. DESHMUKH, *Sentiment analysis on product reviews using machine learning techniques*, in Cognitive Informatics and Soft Computing: Proceeding of CISC 2017, Springer, 2019, pp. 639–647.
- [11] Z. JIANQIANG, G. XIAOLIN, AND Z. XUEJUN, *Deep convolution neural networks for twitter sentiment analysis*, IEEE access, 6 (2018), pp. 23253–23260.
- [12] N. JNOUB, F. AL MACHOT, AND W. KLAS, *A domain-independent classification model for sentiment analysis using neural models*, Applied Sciences, 10 (2020), p. 6221.
- [13] S. KARDAKIS, I. PERIKOS, F. GRIVOKOSTOPOULOU, AND I. HATZILYGEROUDIS, *Examining attention mechanisms in deep learning models for sentiment analysis*, Applied Sciences, 11 (2021), p. 3883.
- [14] X. LI, Y. RAO, H. XIE, R. Y. K. LAU, J. YIN, AND F. L. WANG, *Bootstrapping social emotion classification with semantically rich hybrid neural networks*, IEEE Transactions on Affective Computing, 8 (2017), pp. 428–442.

- [15] N. LIU, B. SHEN, Z. ZHANG, Z. ZHANG, AND K. MI, *Attention-based sentiment reasoner for aspect-based sentiment analysis*, Human-centric Computing and Information Sciences, 9 (2019), pp. 1–17.
- [16] M. MAMTESH AND S. MEHLA, *Sentiment analysis of movie reviews using machine learning classifiers*, Int. J. Comput. Appl, 182 (2019), pp. 25–28.
- [17] M. MUJAHID, E. LEE, F. RUSTAM, P. B. WASHINGTON, S. ULLAH, A. A. RESHI, AND I. ASHRAF, *Sentiment analysis and topic modeling on tweets about online education during covid-19*, Applied Sciences, 11 (2021), p. 8438.
- [18] U. NASEEM, I. RAZZAK, M. KHUSHI, P. W. EKLUND, AND J. KIM, *Covidsenti: A large-scale benchmark twitter data set for covid-19 sentiment analysis*, IEEE transactions on computational social systems, 8 (2021), pp. 1003–1015.
- [19] M. PARIMALA, R. SWARNA PRIYA, M. PRAVEEN KUMAR REDDY, C. LAL CHOWDHARY, R. KUMAR POLURU, AND S. KHAN, *Spatiotemporal-based sentiment analysis on tweets for risk assessment of event using deep learning approach*, Software: Practice and Experience, 51 (2021), pp. 550–570.
- [20] D.-H. PHAM AND A.-C. LE, *Learning multiple layers of knowledge representation for aspect based sentiment analysis*, Data & Knowledge Engineering, 114 (2018), pp. 26–39.
- [21] A. U. REHMAN, A. K. MALIK, B. RAZA, AND W. ALI, *A hybrid cnn-lstm model for improving accuracy of movie reviews sentiment analysis*, Multimedia Tools and Applications, 78 (2019), pp. 26597–26613.
- [22] F. S. RELUCIO AND T. D. PALAOAG, *Sentiment analysis on educational posts from social media*, in Proceedings of the 9th international conference on E-education, E-business, E-management and E-learning, 2018, pp. 99–102.
- [23] F. RUSTAM, M. KHALID, W. ASLAM, V. RUPAPARA, A. MEHMOOD, AND G. S. CHOI, *A performance comparison of supervised machine learning models for covid-19 tweets sentiment analysis*, Plos one, 16 (2021), p. e0245909.
- [24] J. SAMUEL, G. M. N. ALI, M. M. RAHMAN, E. ESAWI, AND Y. SAMUEL, *Covid-19 public sentiment insights and machine learning for tweets classification*, Information, 11 (2020), p. 314.
- [25] A. SCHMIDT AND M. WIEGAND, *A survey on hate speech detection using natural language processing*, in Proceedings of the fifth international workshop on natural language processing for social media, 2017, pp. 1–10.
- [26] C. SINGH, T. IMAM, S. WIBOWO, AND S. GRANDHI, *A deep learning approach for sentiment analysis of covid-19 reviews*, Applied Sciences, 12 (2022), p. 3709.
- [27] C. SINGH, S. WIBOWO, AND S. GRANDHI, *An integrated non-linear deep learning method for sentiment classification of online reviews*, in The International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery, Springer, 2020, pp. 889–896.
- [28] M. SINGH, A. K. JAKHAR, AND S. PANDEY, *Sentiment analysis on the impact of coronavirus in social life using the bert model*, Social Network Analysis and Mining, 11 (2021), p. 33.
- [29] J. SRIKANTH, A. DAMODARAM, Y. TEEKARAMAN, R. KUPPUSAMY, A. R. THELKHAR, ET AL., *Sentiment analysis on covid-19 twitter data streams using deep belief neural networks*, Computational intelligence and neuroscience, 2022 (2022).
- [30] K. VIJAYAPRABAKARAN AND K. SATHIYAMURTHY, *Towards activation function search for long short-term model network: A differential evolution based approach*, Journal of King Saud University-Computer and Information Sciences, 34 (2022), pp. 2637–2650.
- [31] Y. WANG, M. WANG, W. XU, ET AL., *A sentiment-enhanced hybrid recommender system for movie recommendation: a big data analytics framework*, Wireless Communications and Mobile Computing, 2018 (2018).

*Edited by:* Sathishkumar V E

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

*Received:* Mar 9, 2023

*Accepted:* May 2, 2023



## DEEP LEARNING-BASED EDUCATION DECISION SUPPORT SYSTEM FOR STUDENT E-LEARNING PERFORMANCE PREDICTION

**Abstract.** Information Technology (IT) and its advancements change the education environment. Conventional classroom education has been transformed into a modernized form. Education field decision-makers are always searching for new technologies that provide fast solutions to support Education Decision Support Systems (EDSS). There is a significant need for an effective decision support system to utilize student data which helps the university in making the right decisions. The Electronic learning system (e-learning) provides a live forum for faculties and students to connect with learning portals and virtually execute educational activities. Even though these modern approaches support the education system, active student participation still needs to be improved. Moreover, accurately measuring student performance using collected attributes remains difficult for parents and teachers. Therefore, this paper seeks to understand and predict student performance using effective data processing and a deep learning-based decision model. The implementation of EDSS starts with data preprocessing, Extraction-Transformation-Load (ETL), a data mart area to store the extracted data with Online Analytical Processing (OLAP) processing, and decision-making using Deep Graph Convolutional Neural Network (DGCNN). The statistical evaluation is based on the student dataset from the Kaggle repository. The analyzed results depict that the proposed EDSS model on an independent data mart with efficient decision support and OLAP provides a better platform to make academic decisions and help educators to make necessary decisions notified to the students.

**Key words:** data mart, decision support system, deep learning, e-learning, ETL, OLAP

**1. Introduction.** A decision support system is challenging in present-day applications like education, recommendation systems, etc. They provide a framework for decision-making that enables users to analyze and evaluate information from multiple sources, models, and databases to help them make informed decisions. After the pandemic, most learning platforms have developed online services to provide practical learning opportunities to students. Online courses are offered through e-learning (electronic or online learning) to interested learners who can learn from their particular places. The shift from traditional education to e-learning has resulted in several prospects, objectives, and difficulties for educational institutions, instructors, and students. Additionally, the quality of content delivered to students continually improves due to various technological advancements in the animation industry [2]. When student populations increase beyond the e-learning platform, multiple challenges exist in measuring student performance, content quality, etc. Major challenges pointed out by researchers are communication between teachers and learners, online assessment of students, technical issues, student impacts on materials and studies, etc..

A learning management system requires practical tools in an e-learning platform to provide a strong communication link between teachers, students, and administrative staff. E-learning tools help to digitize the classroom environment with good communication. Some applications processed through the e-learning platform are digital class teaching, screen sharing, attendance, assessment, and reports [3]. Recent developments in learning systems like high-speed internet and improved networking structures help increase students' e-learning usage. Student performance can be measured with multi-perspective features from the management system. There are several approaches to measuring student performance and reporting instructions daily in online systems. A decision support system helps mine the relevant data features and provides decisions in time. Deep learning, Artificial Intelligence (AI), and machine learning techniques help to achieve the greater performance of the system.

**Background:** In the field of education, advancements in information technology (IT) have transformed conventional classroom education into a modernized form of education. Educational decision-makers are constantly searching for new technologies. The problem being addressed in this research is the need for an effective decision support system to utilize student data and help universities make informed decisions regarding student performance. While electronic learning systems (e-learning) have provided a live forum for faculties and students to connect and virtually execute educational activities, there is still room for improvement in terms

of active student participation and accurately measuring student performance using collected attributes.

The significance of this research lies in its proposed solution, which is a deep learning-based decision model that utilizes effective data processing and decision-making techniques. The implementation of this Education Decision Support System (EDSS) involves data pre-processing, Extraction-Transformation-Load (ETL), a data mart area to store the extracted data with Online Analytical Processing (OLAP), and decision-making using Deep Graph Convolutional Neural Network (DGCNN).

Primary education to higher education faces significant challenges when the database is huge. Prediction and analysis of big data need more advanced deep learning algorithms in a learning management system. This research investigates a deep learning-based decision support system for analyzing students' performance using an e-learning system. Also, the authors aim to improve the e-learning system by giving the best performance to students. The significant contribution of this paper follows:

1. This paper developed an efficient EDSS for an e-learning environment that can improve the performance of the detection process of student performance.
2. This study consists of four phases: data preparation, ETL, data mart, and decision-making. The first process is data preparation, used to transfer the required dataset from the sources. Next, the extracted data is transformed into the required format using a series of transformation models.
3. The staged data is processed using the ETL transformation process and loaded into the data mart for further processing.
4. OLAP operations are executed in the data mart to make the retrieval process smoother, and then the decision-making process is performed using the proposed DGCNN model.
5. The efficiency of the proposed deep learning based EDSS system with ETL is experimented with and compared to conventional methods using the dataset from Kaggle.

This article is further organized as follows: Section 2 provides details of previous work on similar research, Section 3 presents the proposed work, including algorithms, figures, etc., Section 4 contains information on the results and outcomes, and Section 5 consists of the conclusion of the research.

**2. Related Work.** A methodology by Rujirayanyong and Shi [25] was developed to support universities in making admissions decisions using data mining techniques. The study's objective is to predict applicants' academic performance based on pre-admission criteria, such as high school grade average, Scholastic Achievement Admission Test score, and General Aptitude Test score [24]. The study used a dataset of 2,039 students enrolled in a Saudi public university's Computer Science and Information College from 2016 to 2019. The study results indicate that the applicants' early university performance can be predicted before admission based on specific pre-admission criteria. The Scholastic Achievement Admission Test score is the pre-admission criterion most accurately indicates future student performance. The study also found that the Artificial Neural Network technique has an accuracy rate above 79%, making it superior to other classification techniques [14]. Based on these findings, the authors suggest assigning more weight to the Scholastic Achievement Admission Test score in admissions systems.

The study [21, 26] addresses the challenges of constructing reliable Recommender Systems based on sparse, few in quantity, imbalanced, and anonymized data that may have been stored under poor conditions. This is an essential issue in Data Mining, where the quality and quantity of data can significantly affect the accuracy and reliability of predictive models. The study's success in creating a Recommender System using a real-world dataset from a public Spanish university demonstrates the approach's applicability in natural environments.

The proposed research addresses the limitations of current online exam systems, which rely on traditional machine-learning methods that require handcrafted features and cannot learn hierarchical representations of objects from the data [12]. This limitation affects the efficiency and effectiveness of such systems in detecting cheating. The proposed approach utilizes deep learning models that can automatically extract useful features from visual images and speech using convolutional neural networks and statistical methods. The article's author proposes an e-learning system based on performance analysis using ensemble machine learning. Among other machine learning algorithms, random forest achieves the highest prediction rate in measuring student performance.

In Article [1], Al-Qahtani and Alanzi used student advice and monitoring as a dataset to analyze student performance levels in an e-learning platform. Others have used educational outcomes [6] and student achieve-



ment records [7] as sources of performance analysis. In recent years, there has been significant growth in the use of technology in various aspects of life, including education and healthcare. With the development of intelligent decision support systems and AI-based tools, decision-making has become more efficient and effective. Similarly, the use of technology in online examination systems has led to improved monitoring and cheating detection. This literature review aims to provide a summary of recent research in areas such as intelligent decision-making in healthcare, recommender systems for higher education, deep learning-based cheating detection approaches in online examination systems, AI-based learning style prediction in primary education, teaching machine learning in K-12 classrooms, and the use of AI-based online proctoring systems.

The article [18] proposes an interrelated decision-making model for an intelligent decision-support system in healthcare. The model combines fuzzy logic and neural network techniques to improve the accuracy of the decision-making process in healthcare. The work depicted in [5, 16] develops a recommender system to support higher education students in making enrollment decisions. The system uses a combination of collaborative filtering and content-based filtering techniques to provide personalized recommendations to students. Kadoura and Gumaei propose a deep learning-based cheating detection approach for online examination systems [12, 15]. The approach uses a convolutional neural network (CNN) to analyze students' behavior during an exam and detect any suspicious activity.

The study described in [23] aimed to develop an AI-based learning style prediction model for primary education students in online learning. The model utilizes a hybrid deep neural network (DNN) to predict students' learning styles and provide personalized recommendations to improve learning outcomes. In an article [27], Tedre et al. discussed the pedagogical and technological trajectories for teaching machine learning in K-12 classrooms. They emphasized the need for a comprehensive approach that includes curriculum design, teacher training, and the development of appropriate tools and resources. In their study [1], Al-Qahtani and Alanzi conducted a longitudinal cohort study to compare the predictive values of admission criteria for academic achievement among undergraduate students of health and non-health science professions. The study found that high school GPA was both groups' most significant predictor of academic achievement.

Gopane and Kotecha proposed an AI-based online exam monitoring system to improve cheating detection during exams [6]. The system utilizes a combination of image processing and machine learning techniques to detect any suspicious activity that may occur. In Article [7], Gumaei et al. developed a deep learning-based driver distraction identification framework that utilizes a CNN to detect various driving distractions. This framework can help prevent accidents by alerting drivers who may be distracted. Finally, Gumaei et al. in [8] proposed a decision-level fusion method for predicting the health status of Covid-19 patients. The technique combines multiple machine-learning models to improve the accuracy of health prediction for patients with Covid-19.

Jalali and Noorbehbahani evaluated various online proctoring tools and discussed their advantages and limitations [9]. They highlighted the need for a balance between privacy concerns and effective monitoring. Joshi et al. [10] proposed an automatic method for cheating detection in online exams by processing the students' webcam images. The method uses image processing and machine learning techniques to detect suspicious activity during an exam. Kadam et al. [11] reviewed the literature on explainability in multimodal deep neural networks and discussed various approaches for improving the interpretability of deep neural networks. Kaddoura et al. [13] proposed a method for detecting and localizing multiple image splicing using MobileNet v1. The technique uses deep learning for accurate detection. Furthermore, AI-based decision-making systems for exams and university student performance are discussed with intelligent technologies and advanced tools [19, 20, 22, 28].

Some limitations of the above research are, Firstly, the use of AI-based systems for monitoring exams and identifying cheating has raised concerns about privacy invasion and ethical issues. It is essential to balance monitoring students' activities and respecting their privacy rights. Secondly, the proposed AI-based driver distraction identification framework has limitations in accurately detecting all types of driving distractions. For instance, it may not identify distractions that do not involve physical movements, such as daydreaming or inattention. Thirdly, the proposed decision-level fusion method for predicting Covid-19 patients' health status relies on multiple machine learning models, which may lead to increased complexity and computational costs. Additionally, the accuracy of predictions may depend on the quality and quantity of data used to train the



Fig. 3.1: Overview of proposed Education Decision Support System (EDSS) Source: Authors' elaboration

models. Fourthly, while AI-based methods for detecting and localizing image splicing have shown promising results, these methods may face challenges in identifying more advanced forms of image manipulation, such as deepfakes.

Finally, while AI-based decision-making systems have shown a potential to improve university student performance, there is a need to develop explainable AI models that can provide interpretable results and enhance user trust. Moreover, there is a need for rigorous evaluation of these systems' effectiveness and ethical implications before implementing them on a large scale.

**3. Proposed E-learning Decision Support Methodology.** The overview of the proposed EDSS is depicted in Fig 3.1, which consists of four major components: data preparation, ETL, data mart area with OLAP processing, and decision-making using deep learning. The data preparation phase involves collecting and extracting basic data sources. The extracted input data is then transformed using various techniques to balance the data for further processing. The complete processed input data is loaded and staged in the data mart area for storage. The OLAP processing methods generate various access methods for the data stored in the data mart for analysis. Finally, the proposed EDSS model measures and predicts student performance.

#### 1. Dataset Preparation and Description

During the data preparation phase, various data formats, including structured, unstructured, and semi-structured, are collected and stored in a data warehouse for further processing and decision-making. The ETL process is applied to enhance the decision-making process of the stored data. The primary objective of this study is to evaluate the student's performance and participation during e-learning sessions. The proposed model is evaluated using the Kalboard 360 dataset from the Kaggle repository [2].

#### 2. ETL process

The execution of the ETL process will reduce the data warehouse development time, storage, and cost. It involves various tasks to manipulate the data to obtain the final results. Below, only two first parts of the process are described, and the Load process follows in subsection 3.1.

##### • Extraction

Extraction is the process of reading and understanding the data from various sources and extracting the necessary parts for further processing. The dataset has been extracted using experience API, consisting of 17 features and 480 instances of multiple characteristics. The attributes are classified into integer and categorical types. It comprises 175 female and 305 male students from various countries [2]. The numerous qualities called 'Visited resources', 'Discussions', 'Announcements', and 'Raise hands' were extracted. These attributes help measure the student's functional performance and participation during the class. The constant response to the questions and discussion forum enhances the prediction process. Based on the features, this data set is used

for classification and prediction purposes and divided into three sections such as demographic, academic, and behavioral. The features such as raising a hand during the lecture, learning satisfaction, response to the survey, and using resources demonstrate learner engagement. Apart from this, the features such as gender, semester, nationality, discussion, and so on were selected. The classification of this data set is based on the label column of three categories: ‘low’, ‘medium’, and ‘high’ to indicate the student performance.

- **Transform**

Once the data is extracted and necessary data has been selected, a series of processes are executed to convert the data meaningfully. It consists of five processes: data cleaning, balancing, normalization, and one hot encoding.

- Data cleaning:** Removes incorrect, missing, and duplicate data from the dataset. This step, also called data scrubbing, ensures the classification model’s accuracy, effectiveness, reliability, and efficiency. The blank and white spaces of the labels are avoided, and the redundant features are removed.
- Data balancing:** Data imbalance is a significant problem in machine learning (ML). The data is rebalanced to improve the validation performance. The rebalance has two types: over-sampling (OS) and under-sampling (US). The US reduces the population of the most represented class, and OS increases the minimum defined people of the course. This paper used the best US model, the random US, and the best OS model, SMOTE (Synthetic Minority Over-Sampling Technique), for balancing. The SMOTE model combines the majority of the data with its closest neighbors.
- Data normalization:** The classifier’s performance also depends on features with average or scaled values in the range [3]. This is not necessary to have all the features evenly distributed. This normalization process ensures the ML performance using the methods such as Yeo Johnson transformation, scaling, and one hot encoding.
  - Yeo Johnson:** It is from the Box-cox transformation and not restricted to positive values. It is mathematically stated as in Eqn 3.1

$$X_i^\rho = \begin{cases} \left( \frac{(X_i+1)^\delta - 1}{\delta} \right) & \text{if } \delta \neq 0, \delta \geq 0 \\ \log(X_i + 1) & \text{if } \delta = 0, \delta \geq 0 \\ \left( \frac{-[(X_i+1)^{2-\delta} - 1]}{(2-\delta)} \right) & \text{if } \delta \neq 0, \delta < 0 \\ -\log(-X_i + 1) & \text{if } \delta = 2, \delta < 0 \end{cases} \quad (3.1)$$

where  $X$  is the input data and  $\delta$  is the real number. If  $\delta=1$  means it denotes the identity transformation [17]. The Yeo-Johnson transformation is a mathematical formula that transforms a given set of input data,  $X$ , into a new set of transformed data,  $X^\rho$ , where  $\rho$  is a parameter that can be varied. The Yeo-Johnson transformation allows for data that includes negative values, a limitation of the Box-Cox transformation. The parameter  $\delta$  can take on any real value, including 0 and negative values, which makes it more flexible than the Box-Cox transformation. When  $\delta$  is equal to 1, the transformation reduces to the identity transformation, which means that the transformed data is the same as the original data.

Overall, the Yeo-Johnson transformation is a valuable tool in data analysis. It can transform skewed data into a more normal distribution, making it easier to analyze and model.

- Scaling:** The imbalance feature scales will degrade the classifier performance. It is necessary to normalize it to an acceptable scale. The continuous data values are changed into the range 0 to 1 using the Min-Max method [4] denoted in Eqn 3.2

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad \forall x \in X \quad (3.2)$$

where  $x_{norm}$  is the normalized result,  $x_{min}$  is the value of 0, and  $x_{max}$  is the value of 1 and the remaining values are in the range of 0 to 1. This process makes the features to

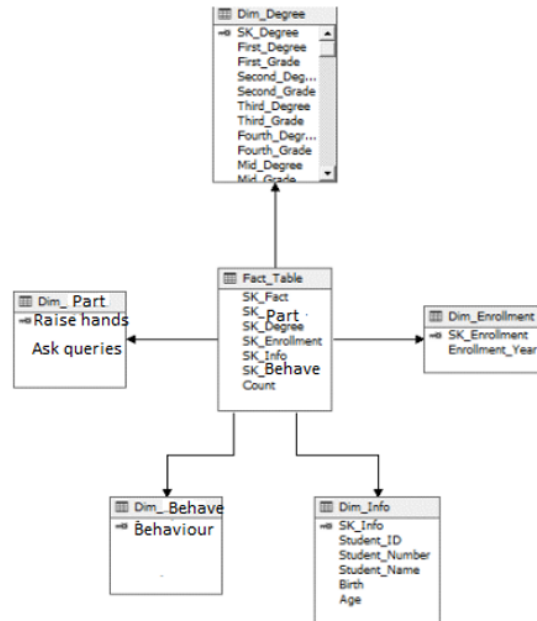


Fig. 3.2: Educational DM (Star Schema) Source: Authors' elaboration

have the same range of values which reduces the training and testing time and overcome the fast convergence rate. This will increase the prediction system's reliability. The log transformation is applied before the normalization to handle the high variance features.

- iii. **One Hot Encoding:** The discrete variables were handled using the label encoders, which encode the categorical data values into an integer. The behavioral attributes of the students are changed into integer values.

**3.1. Data Mart (DM) and OLAP.** The data mart consists of staging and data mart schema tables. The staging area contains the staging table for the transformation process and the final table called the fact table. The data mart schema used in this study is the star schema, where the records are connected to the five-dimension tables using data from the staging table. The staging table is the intermediate area where the storage occurs between the source and the data warehouse. It serves as the temporary storage area that is deleted after being uploaded into the repository. It is used in processes such as data source preparation, cleaning, conversion, extraction, loading, and updating the quality assurance [25], and is referred to by the ETL process. The staging area is prepared with the intention of performing OLAP queries. The data warehouse is implemented using three schemas: star, fact constellation, and snowflake. The star schema is famous for its simplicity, and it consists of a central table (fact table) that concatenates the keys to the dimensions and measurements [24]. Fig 3.2 illustrates the star schema for the Educational DM.

The DM schema is built using SSMS and it consists of five dimensions such as degree, participation, information, enrollment, behavior and degree. The fact tables have five keys which are used to concatenate the dimensional tables. The dimensions are used as the key factor for OLAP process and its responses. The OLAP operations including Roll\_Up, Drill\_Down, Dice, Pivot, and Slice were executed on these dimension tables. Fast responses for the OLAP queries are executed using the star schema. The overall process is stated in Fig 3.3.

**3.2. Decision-Making using DGCNN.** The issue with traditional CNNs is that if the input is not prominent, the CNN pooling layers lose the feature data, which reduces the network's learning ability. There-

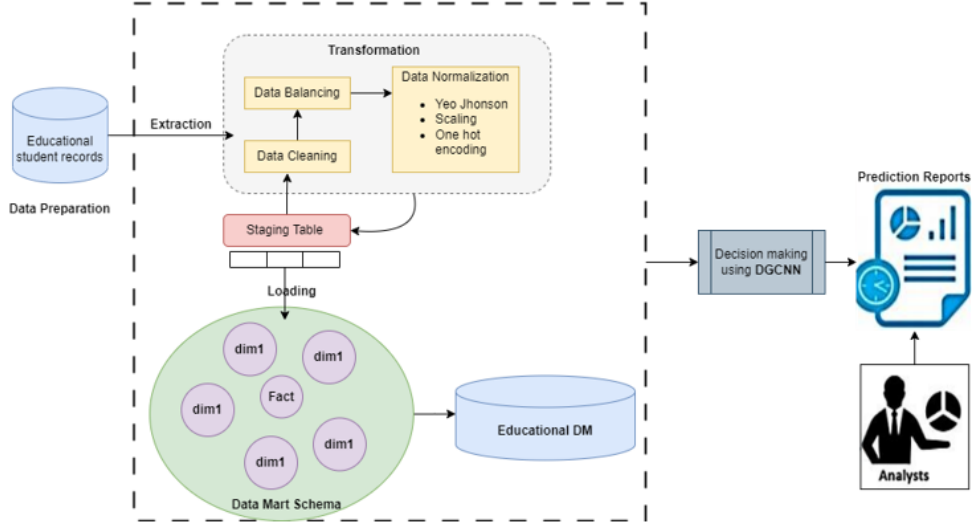


Fig. 3.3: Detailed process of proposing EDSS Source: Authors' elaboration

fore, conventional CNNs do not provide accurate predictions in some instances. Deep graph CNNs are deep learning models that provide an accurate classification for data mining, image processing, and cloud environment problems. They are represented as graphs and process complex data that is not processed by conventional CNNs. The nodes of the graph are connected hierarchically in a spatial area called convolution learning. The graph convolution is constructed with the spectral domain and inverse Fourier transform. The spatial domain is represented by an adjacency matrix  $A \in \mathbb{R}^{(N \times N)}$ , and the spatial relationship between the central node and the neighbor node is determined using spectral similarity. If the node  $X_i$  is adjacent to  $X_j$  then its edges have the weight as  $X_{ij} \in A$ . The node adjacency is represented in Eqn 3.3.

$$X_{ij} = \begin{cases} 0 & \text{if } X_i \text{ is not adjacent to } X_j \\ \exp\left(-\frac{\|X_i - X_j\|}{\sigma_w^2}\right) & \text{otherwise} \end{cases} \quad (3.3)$$

where  $\sigma_w^2$  is the weight range. The DGCNN is declared as  $G(X, A)$  and the classification result is  $Y$  of the graph nodes. This process has three sub-phases such as: The characteristic of each node is extracted. From the local structure of its neighbor, these details are collected. To increase the ability of the network, nonlinear transformation is carried from the previous data node. The graph convolution process aggregates the neighbor node data into a new form which is represented in Eqn 3.4

$$Y = g(f, A) \quad (3.4)$$

$$f^{(i+1)} = \sigma\left(L_g f^{(i)} w^{(i)}\right), L_g = \tilde{D}^{-\frac{1}{2}} \tilde{A}^{-\frac{1}{2}} \quad (3.5)$$

Where,  $f^{(i)}$  is the extracted feature of the layer  $i$  from ETL process,  $\sigma$  is the activation function,  $w$  is the learnable parameter,  $L_g$  is the Laplacian matrix,  $\tilde{D}$  is the degree matrix and  $\tilde{A}$  denotes  $A+I$  and  $I$  denotes Identity matrix. During the training process, the network loss is rescued by adjusting the learnable parameter  $w$  using Eqn 3.6

$$O = \sum_{i=1}^n \text{loss}(g(X_i^g, A), g_i) \quad (3.6)$$

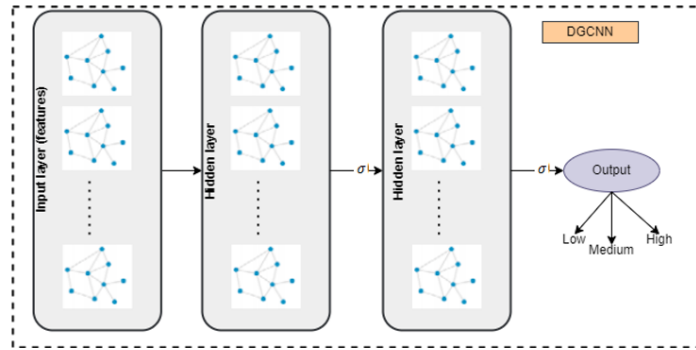


Fig. 3.4: Structure of DGCNN (Source: Authors' elaboration)

where  $g_i$  is the sample ground truth and  $n$  is the number of features. Once the training is over, the feature from the graph and the matrix  $A$  has been used to classify the nodes. The DGCNN structure is shown in Fig 3.4.

The steps to execute the DGCNN are stated as follows:

Input: Dataset  $D$  with features  $f$

Output: Classification results of student performance

Step 1: Collect features from the ETL process

Step 2: Construct the graph nodes called  $V$  and feature  $f$  are used to initialize the graph nodes as  $f^G = [f_1^G, f_2^G, \dots, f_k^G]^T \in R^{k \times D}$ .

Step 3: Construct the graph edges called  $E$  using the first order adjacency relationship by executing Eqn 3.3.

Step 4: Compute the edges  $w$  and form Eqn 3.4

Step 5: Return the prediction results.

**4. Results and Discussions.** The proposed EDSS model has been experimented using the Kallboard dataset. This model is applied using 10 fold cross validation to obtain the optimal accuracy of the learning model. It consists of two parts such as training and testing. In the cross validation, the data set is partitioned into  $k$  folds where each iterations, one subset from testing or training has been used for evaluation. The proposed model is implemented using Python 3.4. from Scikit library using the evaluation metrics such as accuracy, recall, precision and F1 measure.

**4.1. Evaluation Metrics.** The data mining models performance is measured using the factors such as True positive, false positive, true negative and false negative. Based on these factors, the model accuracy, recall, precision and F1score are measured using the following equations.

$$Accuracy = \frac{(Tp + Tn)}{(Tp + Tn + Fp + Fn)} \quad (4.1)$$

$$Recall = \frac{Tp}{(Tp + Fn)} \quad (4.2)$$

$$Precision = \frac{Tp}{(Tp + Fp)} \quad (4.3)$$

$$F1measure = \frac{2 * recall * precision}{(recall + precision)} \quad (4.4)$$

Table 4.1: Proposed DL based EDSS model performance

Class	Accuracy	Recall	Precision	F1measure
Low	0.9348	0.9234	0.9317	0.9145
Medium	0.9472	0.9345	0.9215	0.9176
High	0.9572	0.9478	0.9522	0.9473

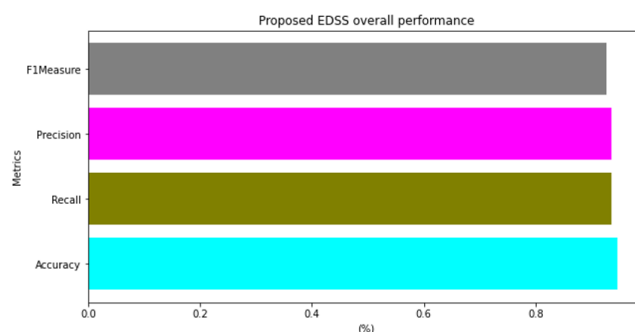


Fig. 4.1: Proposed model overall performance Source: Authors' elaboration

**4.2. Result Discussions.** The performance of the proposed model to classify the student performance under three categories such as low, medium and high is shown in Table 4.1. The overall performance of the proposed model is illustrated in Fig 4.1. As an average of all the classes, the proposed model secured the improved Accuracy of 94.6%, Recall of 93.52%, Precision of 93.51% and F1measure of 92.6%.

The efficiency of the proposed student performance prediction system is compared with the existing decision support system such as C4.5 [14], Artificial neural network [21] and Ensemble ML models [26]. The comparative results are shown in Table 4.2.

From this table, the average performance of proposed and existing approaches are computed and illustrated in Fig 4.2. Compared to the existing approaches, the proposed model secured the improved accuracy of 95%, recall of 94%, precision of 94% and f1measure of 93%. Whereas, the existing DSS model called C4.5 secured the accuracy, recall, precision and f1measure of 88%, 87%, 88% and 87% respectively. The Artificial Neural Networks model secured the accuracy, recall, precision and f1measure of 83%, 82%, 82% and 82% respectively and the Ensemble based DSS secured the accuracy, recall, precision and f1measure of 83%, 82%, 81% and 82% sequentially. The comparison in terms of ROC (Receiver Operator Characteristics) is shown in Fig 4.3 where it proves the efficiency of the proposed model on classifying the student performance as low, medium and high with improved ROC of 0.9 than other approaches. Hence, the proposed model is superior to other approaches on the prediction of student performance.

The average detection time of the student performance is compared, and it is illustrated in Fig 4.4. Due to the implementation of efficient ETL and OLAP operations, the data is stored in the data mart efficiently. The retrieval process is smoother because of this which will reduce the detection rate of the proposed model. Compared to the existing approaches, the proposed model secured reduced detection rate which improves the system integrity, and the required results are displayed in fraction of seconds. The proposed model obtained the reduced detection time of 0.034seconds than other approaches such as C4.5 (1.34s), ANN (1.21s) and Ensemble MLmodel (0.98s) respectively.

The main aim of this study is to develop a new model that can be useful to predict the student performance of e-learning from online series data. The deep learning and efficient ETL process can effectively predict the performance of the students in a minimum detection time which will be helpful for their growth. Compared to the existing DSS models, the proposed model obtained the improved performance from 8 to 9.2% of accuracy.

Table 4.2: Comparative analysis of proposed and existing DSS

Models	Class labels	Accuracy	Recall	Precision	F1 Measure
C4.5	Low	0.8837	0.8729	0.8812	0.8777
	Medium	0.8729	0.8673	0.8714	0.8713
	High	0.8782	0.8711	0.8725	0.8673
ANN	Low	0.8627	0.8536	0.8614	0.8625
	Medium	0.7672	0.7727	0.7525	0.7618
	High	0.8534	0.8253	0.8531	0.8242
Ensemble	Low	0.8684	0.8534	0.8656	0.8684
	Medium	0.7836	0.7727	0.8095	0.7906
	High	0.8351	0.8250	0.7674	0.7951
Proposed ETL+DGCNN	Low	0.9348	0.9234	0.9317	0.9145
	Medium	0.9472	0.9345	0.9215	0.9176
	High	0.9572	0.9478	0.9522	0.9473

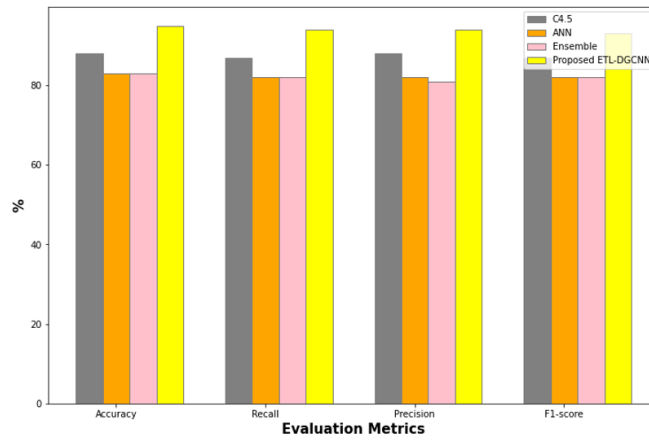


Fig. 4.2: Average performance of proposed and existing EDSS Source: Authors' elaboration

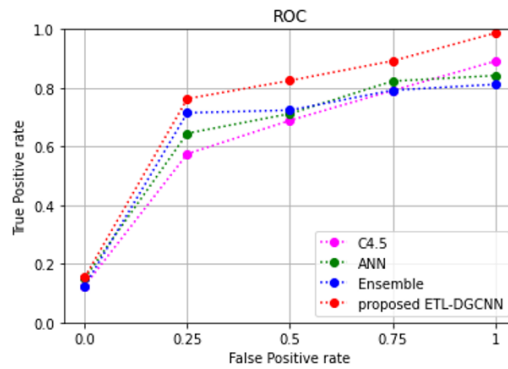


Fig. 4.3: ROC comparison Source: Authors' elaboration



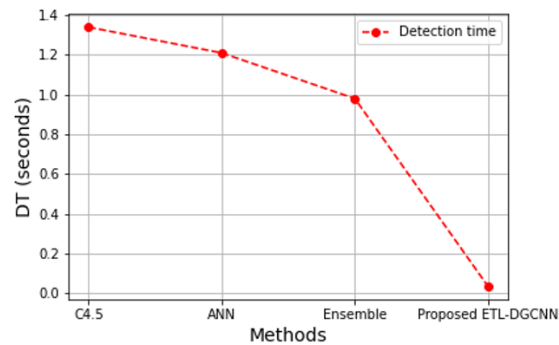


Fig. 4.4: Detection Time comparison Source: Authors' elaboration

Overall, the proposed model results show the accuracy, reliability and efficiency for predicting the performance of the students in the e-learning environment.

**Limitation:** DGCNNs are sensitive to the quality and structure of the input data, which means they may not perform well on datasets with missing or noisy data, or on datasets with complex and irregular structures. DGCNNs are best suited for problems that can be represented as graphs, which limits their applicability to certain domains and types of data.

**5. Conclusion.** This paper proposes the development of an efficient Education Decision Support System for e-learning environments to improve the detection process of student performance. The study comprises four phases: data preparation, ETL, data mart, and decision making. The first phase involves collecting the required dataset from sources and transforming the extracted data into the required format using transformation models. The staged data is then processed using the ETL-transformation process and loaded into the data mart for further processing. In the data mart, OLAP operations are executed to make the retrieval process smoother, and then the decision-making process is executed using the proposed Deep Graph Convolutional Neural Network model. The efficiency of the proposed EDSS system with ETL is experimentally compared with conventional systems using a dataset from Kaggle. The proposed model achieved improved accuracy of 95% with a reduced detection time of 0.034 seconds compared to existing approaches. The timely decisions of the student performance detection can improve their performance in future courses. Hence, the proposed model is suited for the educational industry to make timely decisions with accurate prediction of student performance. In the future, the proposed model will be experimented with an increased number of records in the real-time environment using cloud and Internet of Things.

## REFERENCES

- [1] M. F. AL-QAHTANI AND T. M. ALANZI, *Comparisons of the predictive values of admission criteria for academic achievement among undergraduate students of health and non-health science professions: a longitudinal cohort study*, Psychology Research and Behavior Management, (2018), pp. 1–6.
- [2] A. A. ALSULAMI, A. S. A.-M. AL-GHAMDI, AND M. RAGAB, *Enhancement of e-learning student's performance based on ensemble techniques*, Electronics, 12 (2023), p. 1508.
- [3] J. BROWNLEE, *How to use power transforms for machine learning*, Machine Learning Mastery [Internet], (2020).
- [4] P. DEVAN AND N. KHARE, *An efficient xgboost-dnn-based classification model for network intrusion detection system*, Neural Computing and Applications, 32 (2020), pp. 12499–12514.
- [5] A. J. FERNÁNDEZ-GARCÍA, R. RODRÍGUEZ-ECHEVERRÍA, J. C. PRECIADO, J. M. C. MANZANO, AND F. SÁNCHEZ-FIGUEROA, *Creating a recommender system to support higher education students in the subject enrollment decision*, IEEE Access, 8 (2020), pp. 189069–189088.
- [6] S. GOPANE AND R. KOTECHA, *Enhancing monitoring in online exams using artificial intelligence*, in Proceedings of International Conference on Data Science and Applications: ICDSA 2021, Volume 2, Springer, 2022, pp. 183–193.
- [7] A. GUMAEI, M. AL-RAKHAMI, M. M. HASSAN, A. ALAMRI, M. ALHUSSEIN, M. A. RAZZAQUE, AND G. FORTINO, *A deep learning-based driver distraction identification framework over edge cloud*, Neural Computing and Applications, (2020),

- pp. 1–16.
- [8] A. GUMAEI, W. N. ISMAIL, M. R. HASSAN, M. M. HASSAN, E. MOHAMED, A. ALELAIWI, AND G. FORTINO, *A decision-level fusion method for covid-19 patient health prediction*, Big Data Research, 27 (2022), p. 100287.
  - [9] K. JALALI AND F. NOORBEHBAHANI, *An automatic method for cheating detection in online exams by processing the student's webcam images*, in 3rd conference on electrical and computer engineering technology (E-Tech , Tehran, Iran, 2017, pp. 1–6.
  - [10] G. JOSHI, R. WALAMBE, AND K. KOTECHA, *A review on explainability in multimodal deep neural nets*, IEEE Access, 9 (2021), pp. 59800–59821.
  - [11] K. KADAM, S. AHIRRAO, K. KOTECHA, AND S. SAHU, *Detection and localization of multiple image splicing using mobilenet v1*, IEEE Access, 9 (2021), pp. 162499–162519.
  - [12] S. KADDOURA AND A. GUMAEI, *Towards effective and efficient online exam systems using deep learning-based cheating detection approach*, Intelligent Systems with Applications, 16 (2022), p. 200153.
  - [13] S. KADDOURA, D. E. POPESCU, AND J. D. HEMANTH, *A systematic review on machine learning models for online learning and examination systems*, PeerJ Computer Science, 8 (2022), p. e986.
  - [14] S. LATIF, F. XIANWEN, AND L.-L. WANG, *Intelligent decision support system approach for predicting the performance of students based on three-level machine learning technique*, Journal of Intelligent Systems, 30 (2021), pp. 739–749.
  - [15] S. MAHESWARAN, S. SATHESH, P. PRIYADHARSHINI, AND B. VIVEK, *Identification of artificially ripened fruits using smart phones*, in 2017 international conference on intelligent computing and control (I2C2), IEEE, 2017, pp. 1–6.
  - [16] S. MAHESWARAN, S. SATHESH, G. SARAN, AND B. VIVEK, *Automated coconut tree climber*, in 2017 International Conference on Intelligent Computing and Control (I2C2), IEEE, 2017, pp. 1–6.
  - [17] A. MAHFOUZ, A. ABUHUSSEIN, D. VENUGOPAL, AND S. SHIVA, *Ensemble classifiers for network intrusion detection using a novel network attack dataset*, Future Internet, 12 (2020), p. 180.
  - [18] N. B. MAHIDDIN, Z. A. OTHMAN, A. A. BAKAR, AND N. A. A. RAHIM, *An interrelated decision-making model for an intelligent decision support system in healthcare*, IEEE Access, 10 (2022), pp. 31660–31676.
  - [19] N. MALHOTRA, R. SURI, P. VERMA, AND R. KUMAR, *Smart artificial intelligence based online proctoring system*, in 2022 IEEE Delhi section conference (DELCON), IEEE, 2022, pp. 1–5.
  - [20] M. M. MASUD, K. HAYAWI, S. S. MATHEW, T. MICHAEL, AND M. EL BARACHI, *Smart online exam proctoring assist for cheating detection*, in International conference on advanced data mining and applications, Springer, 2022, pp. 118–132.
  - [21] H. A. MENGASH, *Using data mining techniques to predict student performance to support decision making in university admission systems*, Ieee Access, 8 (2020), pp. 55462–55470.
  - [22] A. S. MILONE, A. M. CORTESE, R. L. BALESTRIERI, AND A. L. PITTENGER, *The impact of proctored online exams on the educational experience*, Currents in Pharmacy Teaching and Learning, 9 (2017), pp. 108–114.
  - [23] B. PARDAMEAN, T. SUPARYANTO, T. W. CENGGORO, D. SUDIGYO, AND A. ANUGRAHANA, *AI-based learning style prediction in online learning for primary education*, IEEE Access, 10 (2022), pp. 35725–35735.
  - [24] P. PONNIAH, *Data warehousing fundamentals for IT professionals*, John Wiley & Sons, 2011.
  - [25] T. RUJIRAYANYONG AND J. J. SHI, *A project-oriented data warehouse for construction*, Automation in Construction, 15 (2006), pp. 800–807.
  - [26] F. SALEEM, Z. ULLAH, B. FAKIEH, AND F. KATEB, *Intelligent decision support system for predicting student's e-learning performance using ensemble machine learning*, Mathematics, 9 (2021), p. 2078.
  - [27] M. TEDRE, T. TOIVONEN, J. KAHILA, H. VARTAINEN, T. VALTONEN, I. JORMANAINEN, AND A. PEARS, *Teaching machine learning in k–12 classroom: Pedagogical and technological trajectories for artificial intelligence education*, IEEE access, 9 (2021), pp. 110558–110572.
  - [28] B. VIVEK, S. MAHESWARAN, P. KEERTHANA, S. SATHESH, S. BRINGERAJ, R. A. SRI, AND S. A. SULTHANA, *Low-cost raspberry pi oscilloscope*, in 2018 international conference on intelligent computing and communication for smart World (I2C2SW), IEEE, 2018, pp. 386–390.

*Edited by:* Sathishkumar V E

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

*Received:* Mar 24, 2023

*Accepted:* May 21, 2023



## FLEXIBLE ENGLISH LEARNING PLATFORM USING COLLABORATIVE CLOUD-FOG-EDGE NETWORKING

**Abstract.** In the modern age, developing practical online learning tools for English language learners is challenging due to existing systems' shortcomings. These systems often need proper instructional design, are well-connected to motivational theories, and have limited infrastructure for data sharing, leading to poor learning outcomes and low motivation. To tackle these issues, a new approach called OAELT has been proposed in this paper. OAELT is an Online Assisted English Learning Tool that uses the Fuzzy Analytical Hierarchy Process (FAHP) and collaborative cloud-fog-edge networking to create a flexible learning design that adapts to the needs and preferences of individual learners. Using the FAHP approach, OAELT provides an improved learning experience by tailoring its design to each learner's unique needs. The collaborative cloud-fog-edge networking approach uses each computing layer's strengths to deliver a personalized and seamless learning experience. OAELT employs adaptive and dynamic approaches within a flexible instructional paradigm to ensure effective instructional design. This paradigm facilitates collective learning data exchange across cloud, fog, and edge computing layers. The effectiveness of OAELT was evaluated using a descriptive statistics approach, which included a five-dimension questionnaire for students covering cognition, emotion, action, cooperation, and literacy. The results demonstrated that OAELT could enhance learning effectiveness and motivation while providing a flexible and seamless learning experience. According to the experimental data of the proposed model, 46.8% of learners often read English magazines and newspapers to improve their flexibility in English learning. Additionally, 50.4% classified and memorized English according to their categories, while 59% of learners often used context to memorize. These findings suggest that the traditional methods for flexible English learning are not adequate, and the average score of the student's methods and strategies is mediocre. However, after using OAELT, some students have been able to use different learning curricular reading. Overall, OAELT's integration of cloud-fog-edge computing with a flexible English learning design can create a more effective and personalized learning system that addresses the challenges of modern learning.

**Key words:** Fuzzy AHP, Cloud-fog-edge collaboration, collaborative interaction, cooperative networks, e-learning, m-grammar learning, adaptive language learning, motivation model, student learning experience.

**1. Introduction.** The demand for effective and efficient learning tools for English language learners has been on the rise in recent years [16]. Traditional language learning approaches have struggled to keep up with modern technologies, resulting in inadequate learning outcomes. To address this issue, modern tools and approaches need to be incorporated to enhance the instructional design and learning process. Self-directed learning and online-assisted language learning tools have become increasingly popular, as they provide learners with independent learning opportunities. Although online-assisted language learning tools have been introduced to develop various language skills, including grammar [16, 12], there are concerns about the effectiveness of existing systems due to a lack of theoretical reference, motivation, and poor instructional design. An adaptive and dynamic approach that caters to learners' needs, as illustrated in Fig 3.1 [16], is required to address these concerns.

The adaptability of the learning system refers to its ability to evaluate learners' actions and guide them in the study process, while the dynamic part refers to the system's flexibility to change its structure according to the demands of the students. However, existing online-assisted English learning applications cannot fully ensure both the adaptive and dynamic attributes of learning, and they have little consideration for motivational factor and cognitive load management [12]. Therefore, there is a need to investigate potential solutions to address these concerns and enhance the effectiveness of online-assisted English learning tools.

Existing instructional tools often suffer from a common problem of providing the same learning materials to all learners, regardless of their proficiency level. Such an approach fails to consider the unique learning characteristics of individual learners, and can lead to boredom or underestimation of the learning material, ultimately resulting in an increased cognitive load. To address this limitation, we propose an adaptive and dynamic learning system, named the Online-Assisted English Learning Tool (OAELT), which takes into account the evolving learning characteristics, motivation, cooperation, and learning performance of individual students.

The aim of the system is to enhance the functionality of existing learning tools by incorporating adaptive and dynamic learning attributes, building upon the networking-assisted learning system that has been previously studied. The ultimate goal is to develop a flexible English learning system that can effectively cater to the individual learning needs of each student.

The proposed Online Assisted English Learning Tool (OAELT) system was designed to provide users a seamless and personalized learning experience. This was achieved by incorporating a flexible English learning design using Fuzzy Analytical Hierarchy Process (FAHP) with collaborative cloud-fog-edge computing. The FAHP approach allowed for a learning design that could adapt to the individual needs and preferences of the learner [14], resulting in a better learning experience. Collaborative cloud-fog-edge networking also played a significant role in the platform's success, leveraging the strengths of each computing layer to provide a seamless and personalized learning experience. The cloud layer provided centralized infrastructure, while the fog and edge layers distributed computing resources closer to the user, enabling faster response times and offline learning experiences [16, 14]. The purpose of this study was to investigate the potential of the proposed OAELT system as a solution to the challenges in English language learning and to demonstrate its effectiveness through quantitative research and collaborative network performance evaluation.

Integrating Fuzzy Analytical Hierarchy Process (FAHP) with collaborative cloud-fog-edge networking has transformed the Online Assisted English Learning Tool (OAELT) system into a more effective, personalized learning solution. This new approach overcomes the challenges of modern learning, providing users with a flexible and seamless learning experience. The system was evaluated using descriptive statistics from a five-dimension questionnaire for students, which assesses cognition, emotion, action, cooperation, and literacy. The evaluation demonstrated that OAELT successfully enhanced learning effectiveness and motivation, making it a promising solution for Online Assisted English Learning.

The main contributions of the paper are as follows:

1. Proposed a novel approach to the Online Assisted English Learning Tool (OAELT) that incorporates a flexible English learning design using Fuzzy Analytical Hierarchy Process (FAHP) and collaborative cloud-fog-edge networking.
2. The proposed OAELT system uses machine learning algorithms to personalize the learning experience, catering to each learner's unique needs and preferences. Customized learning materials and game elements are incorporated to increase learner motivation and engagement.
3. The proposed OAELT system has been shown to be effective in improving its users' learning outcomes and motivation. This was evaluated using a five-dimension questionnaire for students that covers cognition, emotion, action, cooperation, and literacy.
4. The evaluation provides descriptive statistics demonstrating the system's ability to enhance learning effectiveness and motivation while providing a flexible and seamless learning experience.

**2. Literature Review.** Recent research has mainly focused on developing flexible English learning platforms to meet the growing demand for such platforms due to accessibility, convenience, personalization, technological advances, and globalization. However, common limitations in online English language learning have been identified, including a lack of human interaction, limited feedback, motivation, and quality control. Researchers are exploring new technologies and instructional methods such as chatbots, virtual reality, and gamification to address these challenges to provide more interactive and engaging learning experiences. Moreover, efforts are being made to incorporate more opportunities for human interaction and feedback into online learning platforms through features such as video conferencing and live chat with tutors. Additionally, accreditation and quality control measures are being implemented to ensure the quality of online learning content and instruction. Below are some of the studies that are relevant to our work:

According to papers [17, 16] the existing English language learning systems have not been very effective due to three key reasons: firstly, they are not designed based on motivational theoretical principles; secondly, the instructional design is not proper; and thirdly, there is a lack of proper infrastructure for data sharing between students and instructors. To address these issues, this paper proposes MATT: a Mobile-Assisted Tense Tool that uses an m-grammar instructional design and leverages cloud-fog-edge collaborative networking. The Cognitive Theory of Multimedia Learning principles is integrated into MATT to minimize cognitive load, and a motivational model is incorporated to increase motivation and learning effectiveness. To ensure effective

instructional design, the system uses adaptive and dynamic approaches through a flexible instructional paradigm that enables cooperative learning data exchange across the cloud (central unit), fog (regional units), and edge (end devices/learners). Merely utilizing the collaborative cloud-fog-edge technique falls short of meeting these demands, the paper [1] focuses on the security and privacy challenges associated with fog and edge computing. The authors survey the main challenges and explain how these issues can impact the implementation of fog and edge computing. Furthermore, the paper presents various countermeasures to address and mitigate the potential impact of these security challenges.

The author [6] focuses on the controversial nature of human factors constructs and the widespread use of the NASA-TLX (Task Load Index) measurement tool. The author provides a critical perspective on the use of the NASA-TLX, a widely-used tool for measuring perceived mental workload, and highlights some of the challenges and limitations associated with this measure. The paper emphasizes the importance of carefully considering human factors constructs' theoretical foundations and measurement properties and associated measurement tools.

In article [3], the author proposes this article to address the challenges associated with transmitting large amounts of data between edge and cloud computing systems in predictive maintenance (PM) applications. The article discusses the use of federated learning (FL) as a mechanism that allows the creation of a model from distributed data across edge, fog, and cloud layers without violating privacy. However, FL faces challenges in asset management, particularly in PM applications. The article proposes two federated algorithms, Federated Support Vector Machine (FedSVM) and Federated Long-Short Term Memory (FedLSTM) [18], for PM applications, which allow factories at the fog level to maximize the accuracy of their PM models without compromising privacy. The proposed algorithms aim to enable collaborative PM applications that are efficient, accurate, and secure.

This article [19] explores the use of three-dimensional reconstruction of medical images and virtual reality (VR) technology in nursing experiment teaching. The goal is to provide an easier and more effective learning experience for students and simplify teachers' teaching process by incorporating VR technology. The article proposes using VR application technology to create an immersive learning environment that can enhance student understanding and engagement.

Ultimately, the aim is to improve the quality of nursing education through the integration of advanced technologies, [15] suggests that to improve the quality of distance learning, it is important to provide more interactive and diverse content that incorporates multimedia elements like text, images, animations, sounds, and videos. The author proposes a web-based application that combines interactive multimedia and a game-based approach to support distance learning. The aim is to make the learning experience more engaging and enjoyable for students and to provide a more effective way of delivering educational content to learners who are not physically present in a classroom. By using interactive multimedia and a game-based approach, the proposed application could potentially enhance learners' motivation and learning outcomes. The article [17] explains that current systems for teaching grammar often fail to engage and motivate students, which can lead to a higher cognitive load and poorer learning outcomes. To address these issues, the paper proposes a new approach that uses smart communication networks to manage cognitive load and improve the learning experience. By focusing on pedagogically informed instructional design, the proposed system aims to balance the cognitive load better and increase engagement, resulting in better learning outcomes.

Article [7] describes the increasing demand for data processing and optimization due to the billions of data bytes generated at the network edge. As a solution to this challenge, the article proposes the integration of edge computing and artificial intelligence, which is referred to as edge intelligence [8] This work discusses the issue of latency in the communications and processing procedures in a fog network caused by the workload imbalance among IoT devices and base stations (BSs). To address this problem, the author proposes a workload balancing scheme to minimize latency by assigning IoT devices to suitable BSs. This scheme ensures that the workload is distributed evenly among the BSs, which optimizes the processing and communications procedures and minimizes latency [2]. This article discusses how to manage the large amounts of data generated by IoT devices and process it effectively. The proposed solution is to use cloud computing to store and process the data, and IoT can help manage tasks offloaded to the cloud. The article also emphasizes improving application performance by measuring power utilization, makespan, and execution time. The proposed model uses fog

computing to decrease processing time, and its effectiveness is compared to other existing systems.

The article [4] describes a course on autonomous aerial robotics available in the Robotics Academy framework. The study, which is free and open-access, teaches students how to program drones without the need for a physical drone. Students can use their computers to program various types of drones [5] This article reviews 134 research studies on online teaching and learning practices in teacher education. The author identified different practices related to social, cognitive, and teaching presence in online education. These practices are important for ensuring effective online teaching and learning experiences for teachers and students, [9] explains that online learning experiences, which are carefully designed and planned, are distinct from courses offered online in response to a crisis or disaster. Due to the COVID-19 pandemic, many colleges and universities have had to resort to emergency remote teaching to maintain instruction. However, the article emphasizes that this emergency remote teaching should not be equivalent to well-planned online learning experiences. The report highlights the importance of understanding these differences to evaluate the effectiveness of emergency remote teaching during the COVID-19 pandemic.

Article [7] discuss a professional development (PD) program designed for teachers to introduce them to the maker movement and its principles. The goal was to help teachers integrate maker-centered learning into their curriculum. The purpose of this article [11] is to introduce a platform called PaPL, which stands for Paper-based Programming Languages. The report suggests that using paper as a means of interaction is a low-cost and flexible solution that can take advantage of the prevalence of paper in classrooms. The authors aim to make the platform easy to reproduce and accessible for anyone interested in introducing paper-based programming languages in their teaching or learning activities.

Existing online-assisted language learning tools typically follow a pre-designed course structure, which means that learners progress through the content in a linear manner regardless of their individual needs or abilities. This approach can be limiting because learners may not receive the support they need to overcome specific challenges or advance at their own pace. In contrast, OAELT utilizes an adaptive approach that tailors the learning experience to the individual learner's needs and abilities. This is achieved through the use of machine learning algorithms that analyze learner performance and provide personalized feedback and recommendations for further study. Additionally, OAELT's dynamic approach allows for real-time adjustments to the learning experience, enabling learners to receive immediate support when they encounter difficulties or need additional guidance.

By incorporating adaptive and dynamic features, OAELT addresses the limitations of traditional online-assisted language learning tools. It provides learners with a more personalized and responsive learning experience, enabling them to progress at their own pace and receive the support they need to overcome specific challenges. This approach can result in increased engagement, motivation, and ultimately, improved language learning outcomes.

Thus, the above work discusses the advantages and disadvantages of existing teaching modes and how research techniques can be used to create a flexible English learning platform using modern efficient techniques, which will be described below.

**3. OAELT System Model.** The OAELT system model assumes that its components are deployed in the users' devices such as mobile phones, tablets, and laptops. These devices are used to access the system, and the data collected from them is processed using two layers - Fuzzy AHP and Fuzzy Logic.

The Fuzzy AHP (Analytic Hierarchy Process) layer is used to determine the priorities and weights of the different learning elements or parameters. This layer is designed to address the limitations of the traditional AHP method, which assumes that the decision-making criteria are crisp and precise. In contrast, the Fuzzy AHP considers the uncertainty and vagueness associated with decision-making criteria and assigns fuzzy weights to them. This helps in creating a more accurate and flexible decision-making framework.

The Fuzzy Logic layer is used to analyze the collected data and provide personalized recommendations to the learners. This layer uses fuzzy set theory to handle imprecision and uncertainty associated with linguistic variables. It helps in providing learners with more personalized learning experiences based on their individual needs and preferences.

The system also utilizes fog, edge, and cloud layers to control communications between the different components of the system. The fog and edge layers are used to provide early access to the learning contents, which

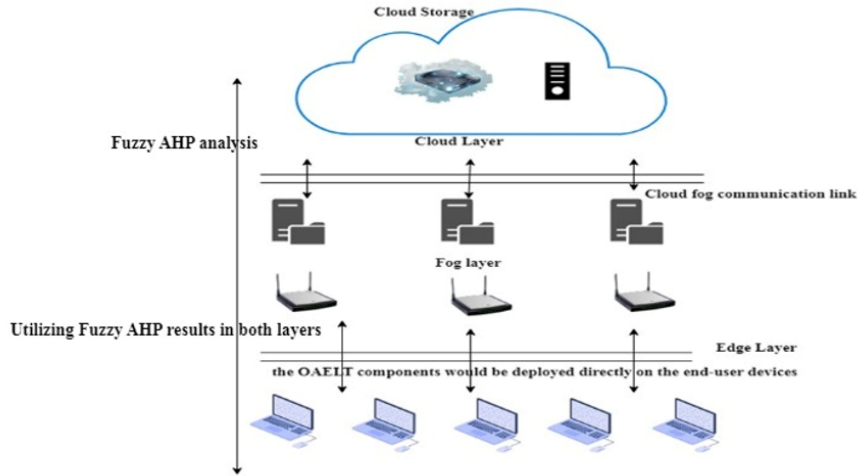


Fig. 3.1: OAELT architecture

helps in reducing the latency and improving the overall performance of the system. The cloud layer is used to store and process large amounts of data and to provide on-demand access to the learning materials

**3.1. FAHP and Collaborative Cloud-Fog-Edge Methodologies.** In this section, we discussed two methods 1. FAHP (Fuzzy Analytical Hierarchy Process 2. Collaborative Cloud-Fog-Edge networking.

**3.2. FAHP.** The Fuzzy Analytical Hierarchy Process (FAHP) is a decision-making methodology that combines fuzzy set theory and the extension principle to provide more accurate results than traditional AHP [14, 13]. The FAHP approach is advantageous when decision-makers are uncertain or vague in their judgments, as it can reduce or remove errors resulting from biases or vagueness.

#### 1. Fuzzy set theory

Fuzzy set theory is a mathematical framework that allows for representing uncertainty and vagueness in decision-making by introducing degrees of membership between 0 and 1. The extension principle is a mathematical principle that calculates the degree of an element's membership in a set. It extends the concept of a function from a subset of a domain to the entire field. In FAHP, fuzzy set theory and the extension principle are applied to the AHP method to enhance its accuracy in decision-making. This section of the paper explains how the FAHP approach uses fuzzy set theory and the extension principle to achieve more reliable decision-making outcomes.

A set of fuzzy numbers can assist decision-making processes, either in triangular form (a, b, c) or interval-valued trapezoidal form (a, b, c, d). Fuzzy numbers represent values that are uncertain or vague and are not easily quantifiable using traditional mathematical methods. In triangular form, a fuzzy number is characterized by three values: the minimum value (a), the most probable value (b), and the maximum value (c). In interval-valued trapezoidal form, a fuzzy number is defined by four values: the left endpoint (a), the left shoulder (b), the right shoulder (c), and the right endpoint (d). These fuzzy numbers are useful in expressing the subjective judgments of decision-makers and enable the integration of imprecision and uncertainty in decision-making procedures [13].

Provides arithmetic operations between two triangular Fuzzy numbers (TFN) [13]. Consider two positives  $m_1$  and  $m_2$  as  $(a_1, b_1, c_1)$  and  $(a_2, b_2, c_2)$  respectively. Fuzzy summation and fuzzy subtraction of two fuzzy numbers are represented as  $\ominus$  and  $\oplus$ .

Assume that

$$\widetilde{m}_1 \oplus \widetilde{m}_2 = (a_1 + a_2, b_1 + b_2, c_1 + c_2) \quad (3.1)$$

$$\widetilde{m}_1 \ominus \widetilde{m}_2 = (a_1 - a_2, b_1 - b_2, c_1 - c_2) \tag{3.2}$$

$$\widetilde{m}_1 \otimes \widetilde{m}_2 = (a_1 a_2, b_1 b_2, c_1 c_2) \tag{3.3}$$

$$\lambda \otimes \widetilde{m}_1 = \lambda a_1, \lambda b_1, \lambda c_1 \tag{3.4}$$

where  $\lambda > 0, \lambda \in R$

$$\widetilde{m}_1^{-1} = \left( \frac{1}{c_1}, \frac{1}{b_1}, \frac{1}{a_1} \right) \tag{3.5}$$

The application of extent analysis principles in comparing two sets of triangular fuzzy numbers (TFNs) [13].

In this context, the objective set is represented by TFNs with membership values given by  $x = \{x_1, x_2, \dots, x_n\}$  and the goal set is represented by TFNs with membership values given by  $u = \{u_1, u_2, \dots, u_n\}$  respectively. To apply extent analysis principles, deriving each object from the sets and performing extent analysis for each goal. This means that for each goal, the degree of membership of each object is evaluated using extent analysis.

$m$  extent analysis values can be derived for each object, where  $m$  represents the number of goals being considered. The values are represented by the formula:

$$m_{g_i}^1, m_{g_i}^1, \dots, m_{g_i}^m, \quad i = 1, 2, \dots, n \tag{3.6}$$

In this formula,  $m_{g_i}^j (j = 1, 2 \dots n)$  represents  $a, ,$  Chang’s extent analysis procedure adopted for fuzzy AHP as follows.

**Algorithm 1: Fuzzy AHP analysis**

Step 1: Obtain the hierarchy structure:

The process of breaking down a problem into smaller parts to make it easier to solve. In this case, the problem is about prioritizing things that are barriers to learning English, and an expert in the field is consulted to help identify the main barriers. Once the barriers are identified, they are organized into a hierarchy, with the most important ones at the top. Then, the barriers are broken down into even smaller parts that need to be considered when prioritizing them. By doing this, it becomes easier to understand and solve the problem effectively.

Step 2: Obtain the pairwise comparison:

Compare the importance of different barriers in English learning, and this can be done by an expert’s opinion. To compare these barriers, experts need to use triangular fuzzy numbers (TFNs) which help to represent the range of possible values for a given parameter. The TFNs have three values, a lower bound, a midpoint, and an upper bound, which represent the expert’s best estimate, range of uncertainty, and imprecision around that estimate. Experts can use TFNs to determine the relationship between two barriers in terms of their relative importance.

Step 3: Evaluation of Fuzzy scores:

$$FS_i = \sum_{i=1}^m m_{g_i}^j \varphi \left[ \sum_{i=1}^n \sum_{j=1}^m m_{g_i}^j \right]^{-1} \tag{3.7}$$

Using Fuzzy summation of TFN  $m$  extent analysis values  $\sum_{j=1}^m m_{g_i}^j$  obtained as

$$\sum_{j=1}^m m_{g_i}^j = \left( \sum_{j=1}^m a_j, \sum_{j=1}^m b_j, \sum_{j=1}^m c_j \right) \tag{3.8}$$



And  $[\sum_{i=1}^n \sum_{j=1}^m m_{gi}^j]^{-1}$  summation of fuzzy  $m_{gi}^j (j = 1, 2, \dots, m)$  values performed to get

$$\sum_m^n \sum_n^m m_{gi}^j = \left( \sum_{i=1}^n a_j, \sum_{i=1}^n b_j, \sum_{i=1}^n c_j \right) \quad (3.9)$$

Inverse vector may be derived as

$$\left[ \sum_{i=1}^n \sum_{j=1}^m m_{gi}^j \right]^{-1} = \left( \frac{1}{\sum_{i=1}^n c_j}, \frac{1}{\sum_{i=1}^n b_j}, \frac{1}{\sum_{i=1}^n a_j} \right) \quad (3.10)$$

Step 4: Obtain the degree of possibility of supremacy

$$m_2 = (a_1, b_1, c_1) \geq m_1 = (a_1, b_1, c_1)$$

$$v = (m_2 \geq m_1) = \sup[\min(\mu_{m_1}(x), \mu_{m_2}(y))] \quad (3.11)$$

Expressed equivalently

$$v = (m_2 \geq m_1) = ht(m_1 \cap m_2) = \mu_{m_2}(d) \quad (3.12)$$

$$\mu_{m_2}(d) = \begin{cases} 1 \\ 0 \\ \frac{a_1 - c_2}{(b_2 - c_2) - (b_1 - c_2)} \text{ if } b_2 \geq b_1, \text{ if } a_1 \geq c_2 \end{cases} \quad (3.13)$$

Step 5: Obtain the degree of possibility for a given convex fuzzy number it is greater than  $K$  convex Fuzzy number  $m_1 (i = 1, 2, \dots, k)$  defined as

$$\begin{aligned} v = (m \geq m_1, m_2, \dots, m_K) &= v[(m \geq m_1) \text{ and } m \geq m_2 \text{ and } \dots \text{ and } (m \geq m_K)] \\ &= \min v(m \geq m_i), i = 1, 2, \dots, K \end{aligned} \quad (3.14)$$

Considering

$$d(a_i) = \min v(s_i \geq s_K) \text{ for } K = 1, 2, \dots, m; K \neq i \quad (3.15)$$

Weight vector derived as  $w = (d(a_1), d(a_2), \dots, d(a_n))^T$  such that  $a_i (i = 1, 2, \dots, n)$  has  $n$  elements

Step 6: Obtain the normalized weight vectors:

$$w = (d(a_1), d(a_2), \dots, d(a_n))^T \quad (3.16)$$

Where  $W$  denotes the non-fuzzy number

Step 7: Compute the overall score

To determine the overall importance of barrier dimensions and factors, priority weightage is calculated using both local weightage and global weightage. Global weightage measures the importance of each barrier dimension or factor across all dimensions and factors, while local weightage measures importance within each group. To obtain the overall score, the global weightages are arranged in descending order for the respective prioritization. This allows for the identification of the most important barrier dimensions and factors, which can then be addressed in order to improve e-learning outcomes.

**3.3. Collaborative Cloud-Fog-Edge (CCFE).** Collaborative Cloud-Fog-Edge (CCFE) networking is an innovative approach to distributed computing that leverages the strengths of cloud computing, fog computing, and edge computing resources to deliver more efficient and flexible services to end-users. This methodology relies on the collaboration between cloud servers, fog servers, and edge devices to provide computing resources, storage, and processing capabilities to end-users which shows in Fig 3.1. Cloud computing, which utilizes remote servers in large data centers, is effective for managing large data volumes and handling complex processing tasks. Meanwhile, fog computing employs intermediate computing nodes positioned closer to the end-users to reduce latency and enhance the response time of services that require real-time processing. Edge computing utilizes local computing resources, such as smartphones, tablets, and IoT devices, to provide localized processing and storage capabilities, and reduce the amount of data transmitted over the network. By combining these diverse computing resources, CCFE networking optimizes computing resources to provide cost-effective services. CCFE networking has various applications, including IoT, big data processing, multimedia services, and flexible and adaptive learning environments in education. By integrating cloud, fog, and edge computing resources, CCFE networking enhances the user experience by providing a more seamless and efficient service delivery, thereby improving the overall quality of service.

In OAELT model the CCFE networking approach provides several advantages to a Flexible English Learning Platform in delivering a more seamless and efficient service. Firstly, it offers cost-effective computing resources for learners by utilizing cloud servers for data management, and fog servers and edge devices for processing and storage. This allows the platform to handle large amounts of data and complex tasks more efficiently, making it possible for learners to access the platform from anywhere using any device, thereby improving their learning experience. Secondly, CCFE networking improves the platform's response time and reduces latency. Intermediate computing nodes can be located closer to learners with fog computing, resulting in faster data processing and better response time. This is particularly beneficial for features such as real-time assessments, where timely feedback is critical for learner progress. Finally, the CCFE approach allows for personalized and adaptive learning experiences. Edge computing provides localized processing and storage capabilities, enabling the platform to collect and analyze data about each learner's preferences, strengths, and weaknesses. Based on this analysis, the platform can tailor learning experiences to each learner, offering customized learning paths, recommendations, and assessments.

**Algorithm 2: Fog-edge- cloud communication**

Step 1: Cloud server executes  
Step 2: Initialize  $w_0$   
Step 3: for each round  $t = 0, 1, \dots, t_g$  do  
    for each factory site  $J \in m$  do  
         $w_{fog,J}^{t+1} \leftarrow fog\ server\ executes(J, w_{cloud}^t)$   
    end  
     $w_{cloud}^t \leftarrow \sum_{J=1}^m a_J w_{fog,J}^{t+1}$   
end  
Step 4: Fog server executes  $(J, w_{fog})$  :  
     $h_J \leftarrow stepsize$   
    for each round  $t = 0, 1, \dots, t_l^J$  do  
        for each edge device  $i \in n$  do  
             $w_{i,J}^{t+1} \leftarrow edge\ device\ update(i, w_{fog}^t)$   
        end  
         $w_{fog,J}^{t+1} \leftarrow \sum_{i=1}^{n_J} \frac{d_i^J}{d^J} w_{i,J}^{t+1}$   
    end  
Step 5:  $w_{fog,J}^{t+1}$  will be return to the cloud server by request  
Step 6: Edge device update  $(J, i, w)$  :  
     $\{d_i\}_{i=1}^n = data\ partition$   
    for each local epoch  $k$  from 1 to  $E$  do  
         $w_{i,J}^{t+1} = w_{i,J}^t - \eta \nabla f_i^J(w_{i,J}^t)$

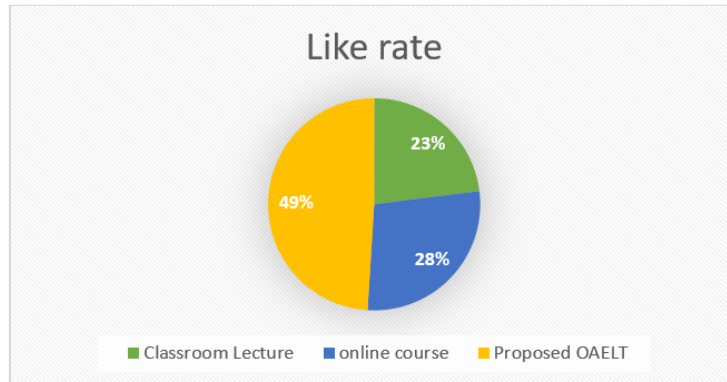


Fig. 4.1: Favourite Learning mode

end

Return  $w_{i,j}$  to  $j$ th fog server

$f_i^J(w_{i,j}^t)$  denotes the loss function for predicting the correctness of the student's English response. We can train the model collaboratively using the cloud, fog, and edge devices to improve the prediction accuracy while minimizing the communication overhead. The partitioned data can be the students' responses and related metadata, such as their English proficiency level, age, and learning history.

#### 4. Results and experiments.

**4.1. Dataset.** Based on the study described in [10], our proposed model involved an experimental setup that included a SPOC course offered on the MOOC platform of Chinese universities. The study was conducted with a total of 120 participants, including 1 lecturer, 1 postgraduate assistant, and 120 students. The students were from two classes of the same grade at Sichuan Normal University and were divided into an experimental group and a control group, each consisting of 60 students. During the preparation stage of the experiment, 120 questionnaires were distributed to investigate the participants' attitude towards online learning, online learning experience, and willingness to serve as a scholar. All 120 questionnaires were completed and returned, and were considered valid for analysis. Favourite mode of learning is shown in Figure 4.1.

Based on the questionnaire results presented in Figure 4.1, it is evident that a significant proportion of students in the experimental group (49%) preferred a learning mode that combined course teaching with online learning. This finding indicates that the majority of students preferred a blended learning model that integrates both traditional classroom teaching and online learning, rather than a single learning model. This suggests that there is a growing demand for a teaching model that combines the benefits of both traditional and online learning methods. In response to this demand, the development of SPOC (Small Private Online Course) courses was initiated. These courses aim to leverage the benefits of both traditional classroom teaching and online learning, providing students with a flexible and personalized learning experience. With the help of the OAELE online assisted English learning tool, which incorporates techniques such as Fuzzy AHP and collaborative cloud-fog-edge networking, SPOC courses can be tailored to suit the individual learning needs of each student. The platform provides a collaborative and interactive learning environment, where students can engage with course materials and peers, at any time and from any location.

In this section, we developed a questionnaire to investigate the cognitive state, emotional state, motor skills, interaction and cooperation, and information technology literacy of the students in both the experimental and control groups, in line with the goals and evaluation system of the OAELE model that incorporates Fuzzy AHP and collaborative cloud-fog-edge networking. The questionnaire was distributed to the students prior to the experiment, and the collected data was organized for better clarity. Descriptive statistics for the five dimensions are presented in Figure 4.2 and 4.3, offering a clear comparison and analysis of the results between the two

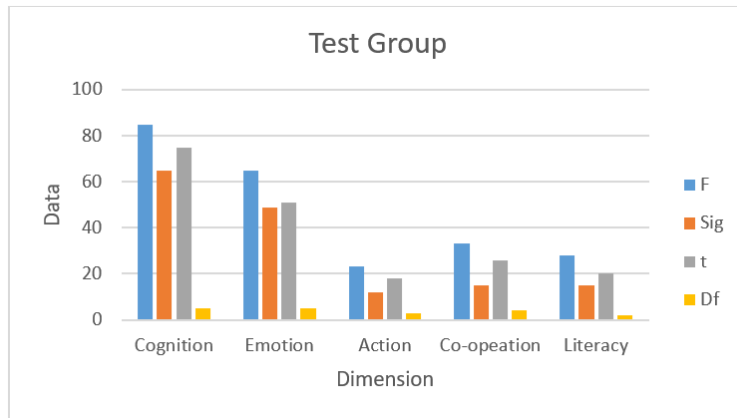


Fig. 4.2: Descriptive statistics for the five dimensions (Test Group)

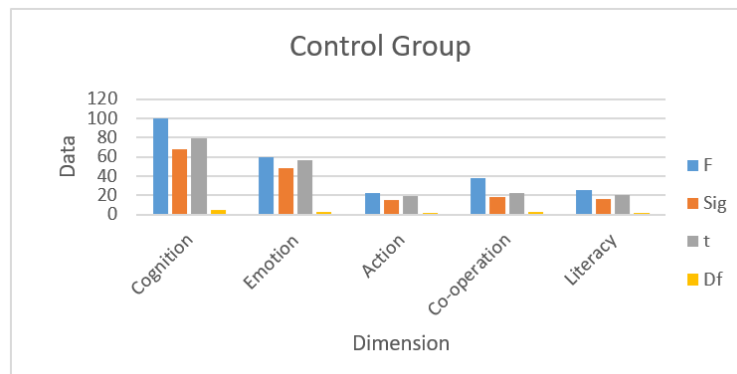


Fig. 4.3: Descriptive statistics for the five dimensions (Control Group)

groups. The F-value (F) , significance (Sig), Test score (t), differences (Df) are used in the analysis features.

The descriptive statistics of questionnaire completed by the students after the experiment were analyzed using OAELT, which incorporates techniques such as Fuzzy AHP and collaborative cloud-fog-edge networking. As depicted in the figure 4.2 and 4.3, the average values of the five dimensions of the students in the experimental group and the students in the control group were found to be similar. This suggests that there is no significant difference in the levels of cognitive, emotional, motor skills, interaction and cooperation, and information literacy between the two groups. To further validate this finding, an independent sample t-test was conducted using SPSS software on the questionnaire data of the experimental group and the control group. The use of OAELT, which incorporates advanced statistical analysis techniques and collaborative learning tools, provided a more accurate and reliable assessment between the two groups. The results of the t-test would help to confirm whether the differences observed in the descriptive statistics of the two groups were statistically significant, further validating the effectiveness of OAELT in facilitating flexible and personalized English learning which is shown in the Table 4.1.

Table 4.1 describing the result of a sample t-test that was conducted to compare two groups of students in terms of several different factors: cognition, emotion, motor skills, cooperation, and literacy. The phrase "no significant difference" means that the differences between the two groups in each of these factors were not statistically significant. In other words, the results of the t-test did not show that the differences between the

Table 4.1: Independent sample t-test before experiment

		F	Sig	t	Df
Cognition	Assuming Equal Variance	4.06	0.03	0.096	117
	Assuming Unequal Variance			0.096	104.78
Emotion	Assuming Equal Variance	1.78	0.28	0.21	117
	Assuming Unequal Variance			0.21	116.08
Action	Assuming Equal Variance	0.50	0.49	0.30	117
	Assuming Unequal Variance			0.30	116.67
Co-operative	Assuming Equal Variance	0.081	0.78	0.64	117
	Assuming Unequal Variance			0.64	116.95
Literacy	Assuming Equal Variance	1.23	0.28	0.25	117
	Assuming Unequal Variance			0.25	115.41

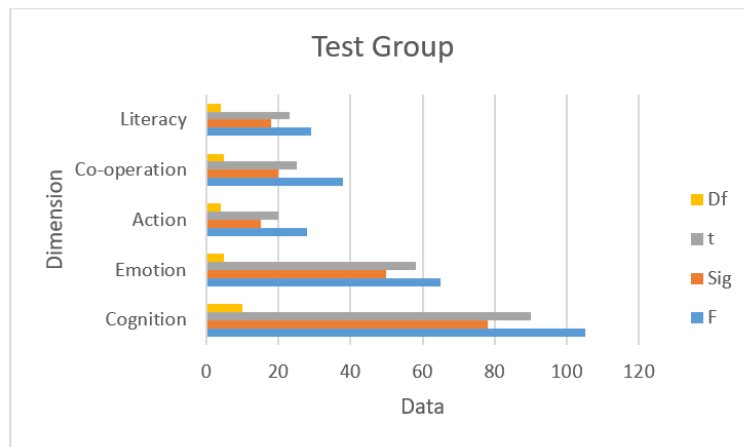


Fig. 4.4: Descriptive statistics of the five dimensions questionnaire for students after the experiment (Test group)

means of the two groups were large enough to rule out the possibility that they were due to random chance. This finding is supported by the data presented in Table 4.1, which presumably shows the means and standard deviations of each group for each of the factors being compared. If the means of the two groups are very similar and the standard deviations are relatively small, this would suggest that there is not a significant difference between the two groups in terms of these factors. The t-test result evaluated using

$$t(df) = t - \text{static}, p = \text{significant value}$$

Using the OAELT, we can conduct experiments to compare and analyze the learning processes and effects of two groups of students. Before the experiment begins, we can conduct an initial assessment or pre-test, which would help us to identify the areas of cognition, emotion, motor skills, interaction and cooperation, and information literacy where the two groups are comparable.

After the experiment, we can use various methods, including experimental observation, written tests, and questionnaire surveys, to analyze the learning processes and effects of the two groups. This surveys would assess the students' cognition, emotion, interaction and cooperation, cooperative learning ability, and information technology literacy. We can use the Fuzzy AHP technique to analyze the data collected from the survey and identify any significant differences between the two groups based on the five dimensions measured by the questionnaire. To analyze the thinking structure of the students, we can use the written test, and the

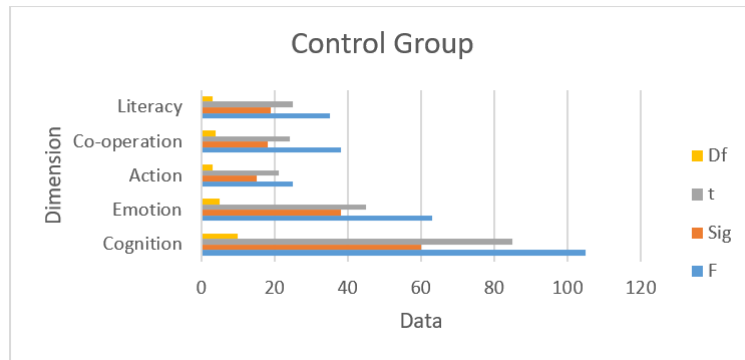


Fig. 4.5: Descriptive statistics of the five dimensions questionnaire for students after the experiment (Control group)

Table 4.2: Independent sample t-test after experiment

		F	Sig	t	Df
Cognition	Assuming Equal Variance	0.67	0.29	6.24	117
	Assuming Unequal Variance			6.24	113.85
Emotion	Assuming Equal Variance	3.74	0.05	5.41	117
	Assuming Unequal Variance			5.41	109.59
Action	Assuming Equal Variance	0.12	0.55	1.82	117
	Assuming Unequal Variance			1.82	116.23
Co-operative	Assuming Equal Variance	0.42	0.38	4.89	117
	Assuming Unequal Variance			4.89	114.71
Literacy	Assuming Equal Variance	4.37	0.02	4.11	117
	Assuming Unequal Variance			4.11	111.85

SOLO layered method can be employed for this purpose. Finally, we can use the collaborative cloud-fog-edge networking technique to create a flexible learning platform that can be accessed from anywhere, at any time, and from any device. With this platform, students can learn English at their own pace and in a way that suits their individual learning styles which is shown in Fig 4.4 and 4.5.

Table 4.2 describing the findings of a study where two groups of students were compared - a "pilot group" and a "control group". According to "Table 4.2", which presumably presents statistical data on the two groups, the pilot group was found to be "substantially significantly dissimilar" to the control group in four dimensions: cognitive, emotional, interaction and cooperation, and information literacy. This means that the two groups were quite different from each other in these areas. More specifically, the students in the pilot group were found to have better cognitive and emotional abilities than those in the control group. They were also found to be better at interacting and cooperating with others during learning activities. However, there was no significant difference between the two groups in terms of motor skills  $P > 0.05$ , which means that both groups performed similarly in this area. To sum up, the students in the pilot group appeared to be stronger in several key areas related to learning than the control group.

**4.2. Experiment of Flexible English learning using OAELT.** In this study, a comparative experiment method was conducted to evaluate the effectiveness of autonomous learning for students in the experimental and control groups. The study utilized a combination of questionnaires, interviews, and subjective test papers to gather data on learners' needs, abilities, and comprehension in English language learning. Advanced techniques such as Fuzzy AHP and collaborative cloud-fog-edge networking were used to analyze the results of

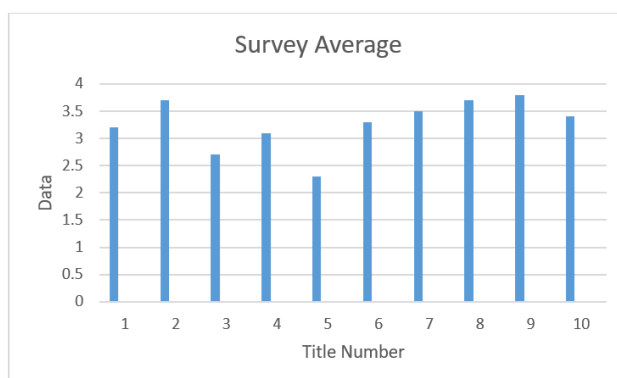


Fig. 4.6: Survey results of student's motivation and interest

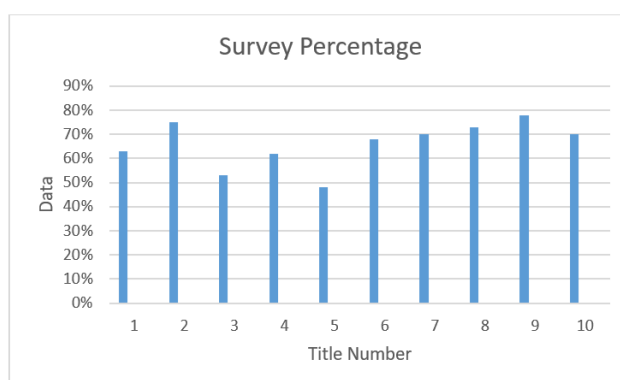


Fig. 4.7: Survey results of student's motivation and interest

the survey and test. OAELT, a flexible English learning platform, was used to support the study and provide personalized learning experiences for the participants. The study's findings, including students' motivation and interest in English language learning, are presented in Figure 4.6 and 4.7. The study's outcomes highlight the importance of personalized support for learners and demonstrate the effectiveness of flexible and personalized learning platforms like OAELT in optimizing learning outcomes.

Based on the OAELT approach, using fuzzy AHP and collaborative cloud-fog-edge for flexible English learning, the data presented in Figure 4.6 and 4.7 suggests that many students may not be fully engaged in autonomous and flexible English learning. The majority of students (63.6%) indicate that they learn English primarily to pass exams, rather than to build their language proficiency and skills. This could be due to external pressure or lack of intrinsic motivation. Additionally, over half of students (53.6%) report that they will not memorize English vocabulary unless assigned recitation homework by their teacher, suggesting a passive and reactive approach to learning. This indicates that they are relying on external cues rather than taking initiative to learn independently. Furthermore, only a minority of learners (45.8%) express confidence in their ability to learn English effectively, indicating that many may lack the self-efficacy and motivation to engage in independent and proactive learning.

Despite these challenges, a significant proportion of students (65.6%) feel that they have gained a lot in learning, indicating that they value learning and its outcomes. The majority of learners (70.8%) also attach importance to learning as a key component of English language acquisition. However, it is important to note that almost three-quarters of students (74.4%) felt the need for training on flexible English learning strategies,

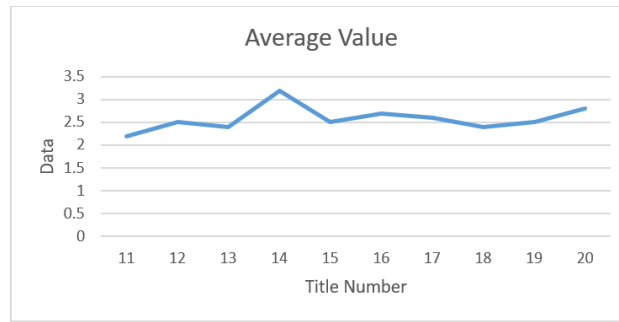


Fig. 4.8: Survey results of students Flexible English learning plans and goals

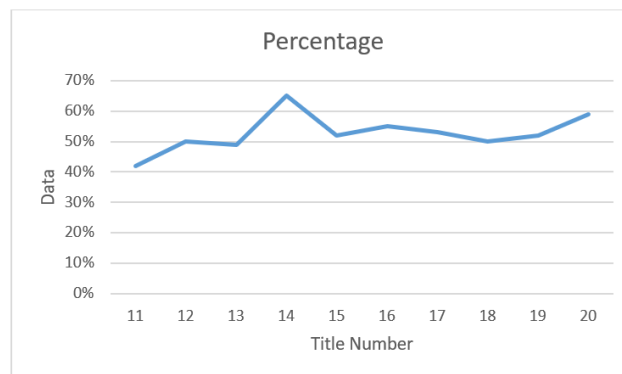


Fig. 4.9: Survey results of students Flexible English learning plans and goals

highlighting the potential benefits of providing students with structured and evidence-based learning resources to support their autonomous learning. By using fuzzy AHP to prioritize these strategies and collaborative cloud-fog-edge to deliver them flexibly, students can access personalized and effective learning materials and support, enabling them to take a more proactive and autonomous approach to their English learning.

Figure 4.8 and 4.9 illustrates the percentage of students who engaged in specific English learning activities. The graph shows that a mere 50.2% of students set goals for learning acquisition, including long-term, mid-term, and short-term objectives. Additionally, 42.8% of students did not establish a personalized learning plan, and 49.2% of students did not allocate a structured study period or prepare a study timetable. Furthermore, only 52% of students developed a periodic review plan for learning, while a significant 59.1% of students reviewed only before the examination.

Here are some potential limitations of the proposed solution and suggestions for future research:

1. **Technology limitations:** While OAELT utilizes cloud-fog-edge networking to overcome some of the limitations of traditional online language learning platforms, there may be technological limitations to its implementation. For example, users may require a high-speed internet connection or advanced hardware to access the platform. Future research could investigate how to mitigate these limitations, such as by developing lightweight applications that can be used on a wider range of devices or providing support to users who may not have access to the necessary technology.
2. **User engagement:** Even with advanced technology, users may not be fully engaged with the platform or motivated to use it. Future research could explore how to improve user engagement with OAELT, such as by incorporating gamification elements, personalization, or social learning features.
3. **Generalizability:** The effectiveness of OAELT may vary across different languages, language proficiency



levels, and learner populations. Future research could investigate how well the platform works across different contexts, as well as identify any potential barriers to implementation or usage.

4. Evaluation metrics: The effectiveness of OAELT may be difficult to measure using traditional evaluation metrics such as language proficiency tests. Future research could explore new evaluation metrics that better capture the unique aspects of the OAELT platform, such as collaboration, information sharing, and technology literacy.
5. Pedagogical approach: While OAELT incorporates collaborative learning and individualized learning approaches, it may not be suitable for all language learning pedagogies or teaching styles. Future research could investigate how OAELT can be adapted to different pedagogical approaches or identify ways to integrate the platform with traditional classroom-based language learning.

**5. Conclusion.** The conclusion of the paper summarizes, that the main points discussed about the challenges of developing effective online learning tools for English language learners due to the shortcomings of existing systems. These shortcomings include a lack of proper instructional design, limited infrastructure for data sharing, and a lack of connection to motivational theories. To address these issues, this paper proposes a new approach called OAELT, an Online Assisted Learning Tool. OAELT uses a combination of Fuzzy Analytical Hierarchy Process (FAHP) and collaborative cloud-fog-edge networking to create a flexible learning design that adapts to the needs and preferences of individual learners. The FAHP approach tailors the learning design to each learner's unique needs, while the cloud-fog-edge networking approach uses the strengths of each computing layer to deliver a personalized and seamless learning experience. This article also highlights that OAELT employs adaptive and dynamic approaches within a flexible instructional paradigm to ensure effective instructional design. This paradigm facilitates collective learning data exchange across cloud, fog, and edge computing layers. The paper evaluates OAELT's effectiveness using descriptive statistics, which include cognition, emotion, action, co-operation, and literacy. The results demonstrate that OAELT can enhance learning effectiveness and motivation while providing a flexible and seamless learning experience. The findings also suggest that OAELT can provide learners with new learning strategies and curricular readings to improve their flexibility in English learning. Overall, the paper concludes that OAELT's innovative approach has promising potential for enhancing online English language learning, providing a more effective and personalized learning system that overcomes the challenges of modern learning.

**6. Funding Support.** This work was supported by Zhejiang Federation of Humanities and Social Sciences Circles (Grant No. 2021N57) and Zhejiang Research Institute of Education Science-Grant No.2023SCG307.

#### REFERENCES

- [1] A. M. ALWAKEEL, *An overview of fog computing and edge computing security and privacy issues*, *Sensors*, 21 (2021), p. 8226.
- [2] D. ARUNKUMAR REDDY AND P. VENKATA KRISHNA, *Feedback-based fuzzy resource management in iot using fog computing*, *Evolutionary Intelligence*, 14 (2021), pp. 669–681.
- [3] A. BEMANI AND N. BJÖRSELL, *Aggregation strategy on federated machine learning algorithm for collaborative predictive maintenance*, *Sensors*, 22 (2022), p. 6252.
- [4] J. M. CAÑAS, D. MARTÍN-MARTÍN, P. ARIAS, J. VEGA, D. ROLDÁN-ÁLVAREZ, L. GARCÍA-PÉREZ, AND J. FERNÁNDEZ-CONDE, *Open-source drone programming course for distance engineering education*, *Electronics*, 9 (2020), p. 2163.
- [5] C. CARRILLO AND M. A. FLORES, *Covid-19 and teacher education: a literature review of online teaching and learning practices*, *European journal of teacher education*, 43 (2020), pp. 466–487.
- [6] J. C. DE WINTER, *Controversy in human factors constructs and the explosive use of the nasa-tlx: a measurement perspective*, *Cognition, technology & work*, 16 (2014), pp. 289–297.
- [7] S. DENG, H. ZHAO, W. FANG, J. YIN, S. DUSTDAR, AND A. Y. ZOMAYA, *Edge intelligence: The confluence of edge computing and artificial intelligence*, *IEEE Internet of Things Journal*, 7 (2020), pp. 7457–7469.
- [8] Q. FAN AND N. ANSARI, *Towards workload balancing in fog computing empowered iot*, *IEEE Transactions on Network Science and Engineering*, 7 (2018), pp. 253–262.
- [9] C. B. HODGES, S. MOORE, B. B. LOCKEE, T. TRUST, AND M. A. BOND, *The difference between emergency remote teaching and online learning*, (2020).
- [10] J. LI, *Adaptive learning model of english vocabulary based on blockchain and deep learning*, *Mobile Information Systems*, 2022 (2022).
- [11] A. MEHROTRA, C. GIANG, N. DURUZ, J. DEDELLEY, A. MUSSATI, M. SKWERES, AND F. MONDADA, *Introducing a paper-based programming language for computing education in classrooms*, in *Proceedings of the 2020 ACM conference on innovation and technology in computer science education*, 2020, pp. 180–186.

- [12] A. MIJUSKOVIC, A. CHIUMENTO, R. BEMTHUIS, A. ALDEA, AND P. HAVINGA, *Resource management techniques for cloud/fog and edge computing: An evaluation framework and classification*, *Sensors*, 21 (2021), p. 1832.
- [13] Q. N. NAVEED, M. R. N. QURESHI, A. O. ALSAYED, A. MUHAMMAD, S. SANOBBER, AND A. SHAH, *Prioritizing barriers of e-learning for effective teaching-learning using fuzzy analytic hierarchy process (fahp)*, in 2017 4th IEEE international conference on engineering technologies and applied sciences (ICETAS), IEEE, 2017, pp. 1–8.
- [14] Q. N. NAVEED, M. R. N. M. QURESHI, A. SHAIKH, A. O. ALSAYED, S. SANOBBER, AND K. MOHIUDDIN, *Evaluating and ranking cloud-based e-learning critical success factors (csfs) using combinatorial approach*, *IEEE Access*, 7 (2019), pp. 157145–157157.
- [15] D. D. PRASETYA, A. P. WIBAWA, AND A. S. AHMAR, *Design of web-based lightweight interactive multimedia for distance learning*, in *Journal of Physics: Conference Series*, vol. 1028, IOP Publishing, 2018, p. 012059.
- [16] N. REFAT, M. A. RAHMAN, A. T. ASYHARI, H. KASSIM, I. F. KURNIAWAN, AND M. RAHMAN, *Matt: a mobile assisted tense tool for flexible m-grammar learning based on cloud-fog-edge collaborative networking*, *IEEE Access*, 8 (2020), pp. 66074–66084.
- [17] N. REFAT, M. A. RAHMAN, A. T. ASYHARI, I. F. KURNIAWAN, M. Z. A. BHUIYAN, AND H. KASSIM, *Interactive learning experience-driven smart communications networks for cognitive load management in grammar learning context*, *IEEE Access*, 7 (2019), pp. 64545–64557.
- [18] F. SAFARA ET AL., *Autism spectrum diagnosis using adaptive learning algorithm for multiple mlp classifier*, *Journal of Intelligent Systems and Internet of Things*, 2 (2021), pp. 33–3.
- [19] M. ZHU, C. LI, S. ZHAO, L. CHEN, X. ZHAO, ET AL., *The role of three-dimensional reconstruction of medical images and virtual reality in nursing experimental teaching*, *Journal of Healthcare Engineering*, 2022 (2022).

*Edited by:* Sathishkumar V E

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

*Received:* Apr 9, 2023

*Accepted:* May 28, 2023



## ANALYSIS, PREDICTION AND CLASSIFICATION OF SKIN CANCER USING ARTIFICIAL INTELLIGENCE – A BRIEF STUDY AND REVIEW

**Abstract.** World Health Organization (WHO) records that skin cancer has vigorously affected people in recent decades. Worldwide, many people are affected by skin cancer, and its affected count will increase yearly. Hence, skin cancer has become a threatening disease. Skin cancer prediction at an earlier time is becoming the higher priority and most challenging task worldwide. A computer-based diagnosis is needed to perform the automatic prognosis of skin cancer. It assists dermatologists in many ways, including the prediction of skin cancer at the earlier stages, easy to diagnose and effective. Nowadays, artificial intelligence-based machine learning approaches have been implemented for an early prediction of cancer in the skin through medical images. This paper is focused on a detailed, comprehensive review of skin cancer analysis, forecast, and algorithmic-based procedures for classifying skin diseases. Moreover, this review paper focused on various stages of algorithm approaches for skin tumor detection like pre-processing data, segmenting data, feature selection, and disease classifier. This detailed review of neoplasm diseases like cancer on the skin is done based on machine and deep learning algorithms to help further research.

**Key words:** Skin cancer, Deep learning, Algorithmic based detection, Artificial intelligence, Machine learning, dermatologist.

**1. Introduction.** Analysis of health images using image processing, segmentation, and disease classification plays an essential part in skin cancer diagnosis. It assists the dermatologist in making proper decisions and earlier diagnosis of disease [48]. The Society of American Cancer reports showed that high mortality rates of cancer in the derma increase the mortality rate of patients by 75%. The highest mortality rate of melanoma increased by 14%. Similarly, based on the cancer report of the American Association in the year 2022, in the US, the mortality rate of tumor patients has steadily declined [65, 57]. Compared to the United States in China, new cancer patient cases have increased daily. Therefore, the mortality rate also increased five times [66].

Due to tremendous changes in the world's environmental situation, we noticed that many people are affected by cancer. Skin cancer causes mainly due to ultraviolet radiation (UV). In recent decades people have been vigorously affected by skin neoplasm. Skin cancer is differentiated into melanoma and nonmelanoma skin tumors. A statistical report by the American Cancer Reports that melanoma-type cancer has a high mortality rate in humans [8, 63, 15, 13]. Skin cancer is diagnosed visually using clinical screening. Then it can be analyzed by applying a biopsy procedure, histopathological images, and the Dermoscopic method [5]. This examination procedure is time-consuming, slow, and painful. Therefore, algorithmic-based diagnosis of skin cancer produces accurate, less expensive, and speedy processes [21].

Heckler et al [29] described implementing artificial intelligence concepts used in the predictive skin cancer analysis and classifying the severity of diseases. Also, this method applies the convolutional neural (CNN) network concept for differentiating skin viruses. Brinker et al. [12] described a CNN for skin cancer classification of patients along with dermatologists for speedy diagnosis. Guha et al. [59] proposed the SVM algorithm with VGGNet for predicting and classifying skin tumor cells. Attia et al. [9], presented an analysis of skin lesion images based on generative adversarial networks with realistic hair simulators.

The process of skin cancer detection using deep learning typically involves several stages. Firstly, a large dataset of annotated skin images is collected, consisting of both cancerous and non-cancerous lesions. These images serve as the training data for the deep learning model. The CNN architecture is then designed and trained on this dataset to learn the intricate patterns and features that distinguish different types of skin cancer. During training, the deep learning model learns to automatically extract relevant features from the skin images, allowing it to identify specific patterns associated with malignant tumors. The model is optimized by iteratively adjusting the network's weights and biases to minimize the error between the predicted and actual labels of the training data.

Once the model is trained, it can be deployed for real-time skin cancer detection. New, unseen skin images

can be fed into the trained model, and it will output a prediction or probability score indicating the likelihood of malignancy. Dermatologists can then utilize these predictions as an aid in their diagnosis, combining the model's output with their clinical expertise for a more accurate assessment. The use of deep learning for skin cancer detection offers several advantages. These models have demonstrated high accuracy rates, often comparable to or even surpassing human dermatologists. They can analyze images rapidly, potentially reducing the time required for diagnosis. Deep learning algorithms can also handle large amounts of data and generalize well to new, unseen cases, making them valuable tools in screening and triage scenarios.

However, there are still challenges to overcome in this field. Deep learning models heavily rely on the availability of large, diverse, and accurately annotated datasets for training. Obtaining such datasets can be time-consuming and resource-intensive. Furthermore, the interpretability of deep learning models remains a concern, as understanding the internal workings and decision-making processes of these complex networks can be challenging.

The main motivation behind this research study is to reduce the death in skin cancer, which is curable when predicted in early stage. Many research works have been implemented in the prediction of cancer in the skin. This article is focused on a detailed, comprehensive review of skin cancer detection and algorithmic-based procedures for classifying skin diseases. As well as reviewed various processing algorithms on image data for the prediction of skin tumor cells. It also implements the various image-processing-based classification algorithms associated with detecting skin lesions and their severity.

The paper is presented as shown: Related work section on skin cancer was discussed in Section 2, and Section 3 is about an overview of skin layers and their diseases. Section 4 describes the types of skin cancer. Section 5 describes various types of algorithmic procedure for skin cancer identification, Section 6 discusses the detection of skin tumor using computer vision processing techniques, Section 7 about description of the dataset, Section 8 discuss challenges in skin cancer prediction, Section 9 discusses the performance of metric measures in the prediction of skin cancer, Section 10 finally summarizes the paper.

**2. Related Work.** Due to the increasing complexities, environmental changes will automatically affect human beings, and dermatological disorders arise become a significant challenging medical issue. Especially recently, skin cancer has become a significant issue and may increase the death rate. For earlier skin cancer prediction, dermatologists expect an automated model of algorithmic-based approaches. Therefore, based on the image processing techniques. A skin cancer diagnosis is a better technique to implement it.

One of the predominant tumor types is skin-related cancers, categorized into melanoma and non-melanoma. High ultraviolet (UV) exposure to sun rays is one of the reasons for skin cancer. The categories of Melanoma cancer develops the pigment cells in the skin (melanocytes). This pigment cell creates changes in your skin color. But it is most commonly found on the chest, back, or men's and women's legs. It can also affect the mouth, anus, or genital area. The reasons behind melanoma are genetic factors, fair skin, unusual moles, less body immunity, tanning beds, and excessive exposure to UV rays [33].

Long Zhang et. al [73] presents that metaheuristic optimization algorithm for skin cancer image classification using a CNN model. This classification algorithm uses a whale optimization model to improve the weight and bias value in the CNN. It is analyzed using the DermIS Digital Database and Dermquest Database datasets. This approach produces better precision values compared with other classification techniques. Amirreza Mahbod et.al [40] propose automation for lesions classification in the skin with the CNN technique for extracting the features of the tumor image and pre-trained the model using various deep models of ResNet-18, AlexNet, and VGG16 for the classification of lesions in skin; it uses an SVM classifier. Hardik Nahata et. al establish the CNN model in diagnosing and detecting tumors in the skin using images. Enhancing robustness in the image classification in skin lesions requires pre-processing data augmentation techniques. For the early diagnosis purpose, this paper implements the Transfer Learning techniques. Francesco Rundo et. al [53] proposed that an early diagnosis of skin neoplasms, or abnormal growths of cells, is essential for effective treatment and management. Non-invasive techniques, such as deep learning CNN, can be helpful. It analysis the skin Lesion using a morphological operator. Table 2.1 shows the survey on the study of skin cancer.

In this literature review, we analyse many researches works on detecting skin cancer. The study analyse major limitation of the conventional methods. Some dataset remains small and achieving results on small datasets are not permanent. The more studies use deep learning models which proves that deep learning achieves

Table 2.1: Skin cancer survey

Author	Algorithm used	Diagnosis of Skin Cancer	Dataset
Fu, Zexian et.al (2022)	The kernel fuzzy C-means (KFCM) is a variant of the fuzzy C algorithm, with an optimized Red Fox Optimization algorithm	melanoma using dermoscopy lesions	ISIC 2020 database
Arun Raj et.al [36] (2022)	CNN, AdaBoosting, Gradient Boosting, and Decision Tree	Melanoma	100 dermoscopicimages
Jojoa Acosta et.al [32] (2021)	Region-based CNN withResNet152	Benign/malignant	ISICdataset
Kim et.al [35] (2021)	Generative Adversarial Networks used as Unsupervised Feature Elimination	Melanoma Classification	ISICdataset
Wang et.al [64] (2021)	Random Forest, k-means,SVM and ResNet	Benign/malignant	122 malignant (i.e., cancerous) and 196 benign skin lesions(non-cancerous).
Sikkandar et.al [67] (2021)	Adaptive NeuroFuzzy (ANFC) classifier and grab cut model.	skin lesion segmentation	ISICdataset
Bumrungkun et.al [14] (2021)	Geometrical templates of gradient vector flow (GVF) and active contours are techniques used for boundary segmentation of skin cancer.	Segmentation of skin lesions(non-cancerous)	227 skin lesion images of skin cancer
Sagar et.al [54] (2020)	ResNet-50 is a CNN architecture with deep transfer.learning	Melanoma Classification of Images	ISIC dataset
Abder-Rahman Ali et.al [7] (2020)	NaiveBayes algorithm, CNN is used to detect cancer.	skin lesion segmentation	307 images from the “ISIC 2018
K Srinidhi et.al [56] (2020)	CNN model using ReLUactivation function	Detection of melanoma skin cancer	ISIC 2019
Nida N et.al [46] (2019)	For clustering the similarity of pixel values by fuzzy C-means clustering with Deep region-based CNN	Lipoma/ fibroma/ sclerosis/ melanoma	ISIC dataset
Albahar [5] (2019)	CNN with 2-layer architecture.	Benign/malignant	The used dataset is ISIC
Bisla, D et.al [11] (2019)	Deep convolutionalGAN	Melanoma	ISIC 2017, ISIC2018, PH2
Vijayalakshmi [62] (2019)	CNN and SVM	Malignant	ISIC dataset
Ali et.al [1] (2019)	Self-attention-basedPGAN	BCC, benign keratosis	ISIC 2018
Mendes, D.B et.al [43] (2018)	CNN with Res-Net 152architecture	Malignant melanoma and BC carcinoma	Two datasets are used. The first data set contains 170 skin images and another dataset contains 1300 skin images
Dorj et.al [22] (2018)	SVM with deep CNN	CC/ SCC/ melanoma	3753 dermoscopicimages
Rezvantalab et.al [51] (2018)	Deep CNN architecture in the version of Inception ResNet v2, Inception v3,DenseNet 201	malignant (i.e., cancerous) benign skin lesion (non-cancerous).	HAM10000 and PH2dataset
Sajid et.al [55] (2018)	Modified KNN	Automated Otsu method of thresholding for segmentation	500 lesion images

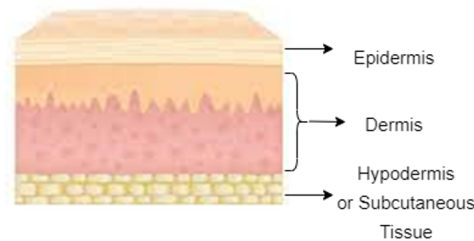


Fig. 3.1: Layers of Skin

promising results on the classifying the cancerous part in the skin. Fuzzy clustering algorithms like KFCM can be complex and challenging to interpret, making it difficult for clinicians to understand the rationale behind the classification decisions. This lack of interpretability could hinder the adoption of the approach in clinical practice. Although modified KNN is traditional classifier, sensitive applications like skin cancer prediction needs strong segmenting and classifying technique.

**3. Overview of Skin Layers and Their Diseases.** This section presents these in structure and discusses the outline of skin diseases. Skin is the human body's largest organ. It is formed up of multiple layers of cells and tissues that protect our parts from the environment and it guards our body against harmful things like extremely high temperatures, sunlight, etc. It consists of the dermis, hypodermis, and epidermis. This structure of the skin provides thermoregulation sensation, and protection [47, 42]. The layers of skin are shown in Figure 3.1.

The observation of Figure 3.1 shows that the outermost layer is the epidermis, which protects our body from the aggressiveness of the environment through the aegis. The dermis layer appears below the epidermis layer, composed of tough tissue, sweat glands, hair follicles, sensors, collagen fibers, and blood vessel receptors. This composition of the dermis layer provides variations in the appearance of the skin. The hypodermis/subcutaneous tissue is the skin's inner layer. This layer interconnects the layer called the dermis with muscles and bones. It protects our body from sweat, cold, and heat and regulates the body temperature [47, 41].

The main causes of skin diseases are environmental factors of temperature, harmful ultraviolet (UV) sunrays, and friction. And also depends on biological factors like allergies, viral infections, and insect bites. Due to genetic factors also, it creates skin diseases. The appearance of skin diseases is air bubbles, different forms of skin tone, hair fall in the skin, non-uniformity of skin, and so on [71]. Skin diseases are mostly because by bacterial, fungi, and viral infections. These infections include erysipelas, trauma-related infections, skin ulcers etc. These skin infections are spread further, and their severity may create health issues like fever and pain or without pain [69]. In some cases, symptoms of skin diseases are rashes; these rashes are used to diagnose certain diseases like rubella, measles, chickenpox, and erythema infectiosum [58].

**4. Skin Cancers Types.** Here we present the various types of skin cancer. Healthy growth cells of melanocytes will grow beyond the control, creating skin tumors. Melanoma is a tumor infection of the derma that begins in the derma cells that produce pigments called melanin. These cells are called melanocytes. It gives colors to eyes, hair, and skin. Melanoma occurs in all body parts, but it commonly occurs in skin that exposes in the sun. Sometimes skin areas not exposed to sunlight can also be affected. For example, hand palm and feet sole. Melanoma skin cancer can be cured early at the detection and diagnosis. Otherwise, it can be spread all over the body [34]. Types of melanomas are 1.Nodular, 2. Spreading superficially 3. Maligna Lentigo 4. Lentiginous in Acral, and 5. Mucosal. The non-melanoma cancer can classify as three types: 1. sebaceous gland carcinoma (SGC), 2. Basel Cell (BCC) and 3. squamous cells scarcinoma (SCC). This non-melanoma type tumor grows slowly that begin in the basal cell, that located in the lower region of the epidermis (skin-outer layer). When compared with melanoma cancer, this nonmelanoma is a curable one. Therefore, it is necessary for early skin cancer diagnosis and treatment [60]. Figure 4.1 shows the various skin tumor types.

Figure 4.1 shows four types of derma tumors Merkel cell, basal cell, Melanoma, and Squamous cell cancer.



Fig. 4.1: Types of Skin Cancer

*Basal cell type carcinoma.* This type of basal cell will appear in the lower epidermis as a round. It is probably in the neck and head and can be identified anywhere on the skin. The main reason this type of cancer is prevalent in areas that have sun exposure, such as the neck, face, and arms. It develops in the body and slowly spreads to the remaining parts of the body. Similarly, 80% of skin cancers belong to this type of skin cancer.

*Squamous cell carcinoma.* The structure of Squamous cell type carcinoma is the form of flat cells called squamous cells. This type of cancer is caused by exposure to the skin in sunlight. It can be shown in the skin as burned or injured by chemicals. Nearly 20 % of skin cancers belong to this tumor type. Similarly, this variant of Squamous cell carcinoma cancer is developed in the lips, outside the mouth, vagina, and anus. This cancer can be affected the remaining parts of the body in the ratio of 20 % to 50 %.

*Merkel cell cancer.* This is a rarely occurring cancer, but it is fast, aggressive growth in the skin. It is produced in the skin because of hormonal changes in the cells under the skin and also in the hair follicles. Usually, this type of skin cancer appears in the head and neck region. It is also called neuroendocrine tumors.

*Melanoma Skin cancer.* The cells affected by melanoma are called melanocytes, in which the epidermis meets the dermis. This cell creates the pigment, and it changes the color of the skin. It is aggressively found in skin cancer.

**5. Analysis of Skin Cancer Based on Various Algorithmic Procedures.** This section explains the analysis of skin cancer using different algorithms. Skin ailments are diagnosed using visual analysis of the appearance of the skin or analysis of clinical screening. For the accurate diagnosis of skin disease, algorithmic-based techniques are implemented. Figure 5.1 shows the algorithmic procedure for skin cancer analysis. Deep learning (DL) and machine learning (ML) algorithmic procedures used to analyze skin cancer diseases developed various approaches like supervised learning and unsupervised learning.

The automatic intelligent model of prediction, detection, and skin cancer diseases classifier is implemented in healthcare development, including dermatology. And it can be used to develop a model for evaluating the skin images from the dataset to analyze skin cancer. For efficient, accurate, and early skin cancer diagnosis, the DL and ML approaches were used. It saves the life of human beings and reduces the health of the patients [20]. The artificial intelligence concept can accomplish skin cancer detection, which is better than human qualities. The deep learning algorithm is used in the dataset of International Skin Image Collaboration (ISIC)-2016 for the analysis of skin diseases like benign or malignant, and it compares 100 skin lesions along with dermatologists [18]. Train the

Dermoscopic dataset with large benign lesions and melanoma images using deep learning algorithm and it uses the InceptionV4. This performance is compared and analyzed by 58 dermatologists [27].

Figure 5.1 shows that based on the algorithmic procedure of Deep (DL) and machine learning (ML) used to analyze skin cancer diseases developed, various approaches like supervised learning and unsupervised learning. For the DL approach, CNN, ANN, and RNN come under the categories of supervised learning approach. Similarly, SOM, BM, and autoencoder are implemented in the unsupervised learning algorithm. In the machine

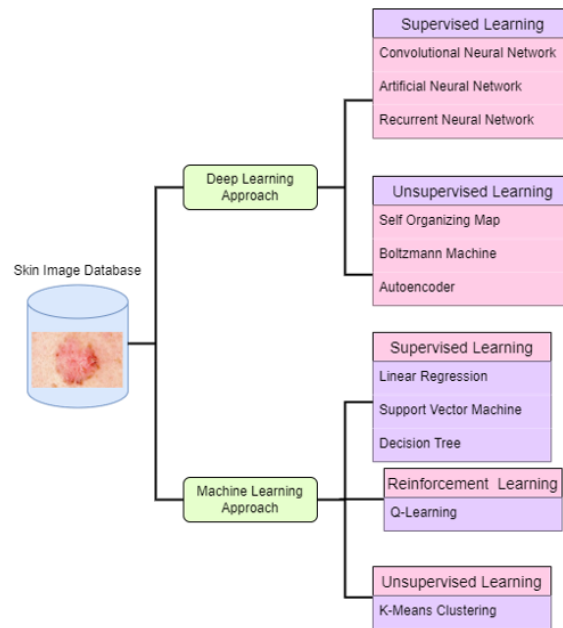


Fig. 5.1: Algorithmic Approaches for the Analysis of Skin Cancer

learning algorithm, LR, SVM, and DT are implemented in supervised learning algorithms. The reinforcement learning algorithm of Q-Learning, KCM for unsupervised learning algorithm is implemented.

*Unsupervised Learning Algorithm.* To make effective decision-making in the skin images analysis using a training dataset this unsupervised learning algorithm is used. It trains the dataset of skin images, using DL and ML for the analysis and classification of skin tumors [50]. Unsupervised DL approaches use the techniques of an iterative procedure, thresholding, and statistical region merging [10]. High-level extraction of features from the image and classifies the input image based on its probability distribution of features. Even though it does not require an extensive training input dataset and for the medical image analysis of skin lesions implement the diffusion of brightness of the medical image which contains multiple peaks. It also has limited capacity for accurately segmenting medical images like skin lesions since their appearance is in various variations like rough skin tone, artifacts, etc. Nowadays, for the study of medical illustrations the DL and ML techniques of Restricted Boltzmann machines (RBM), Deep Boltzmann machines (DBM), Generative adversarial network (GAN) auto-encoders, Deep belief networks (DBN) are used.

Unsupervised learning algorithms play a crucial role in skin cancer detection, particularly in tasks such as clustering, anomaly detection, and feature extraction. Although supervised learning methods dominate the field of skin cancer classification, unsupervised learning techniques offer unique advantages and can complement traditional approaches. Here are a few examples of unsupervised learning algorithms used in skin cancer detection:

1. **Clustering Algorithms:** Clustering algorithms group similar instances together based on the inherent patterns and structures in the data. In the context of skin cancer detection, clustering algorithms can be used to identify different subtypes or clusters of lesions. This can help dermatologists gain insights into the heterogeneity of skin cancer and potentially uncover new patterns or correlations. Popular clustering algorithms include k-means clustering, hierarchical clustering, and density-based clustering.
2. **Dimensionality Reduction Techniques:** Unsupervised dimensionality reduction techniques are employed to reduce the complexity and dimensionality of the input data while preserving the most relevant information. Principal Component Analysis (PCA) and t-Distributed Stochastic Neighbor Embedding (t-SNE) are commonly used techniques in skin cancer detection. By transforming the high-dimensional



image data into a lower-dimensional space, these algorithms facilitate visualization and analysis of the data, aiding in the identification of meaningful patterns and features.

3. **Anomaly Detection:** Anomaly detection algorithms identify instances that significantly deviate from the expected patterns or behavior in a dataset. In the context of skin cancer detection, these algorithms can be used to identify rare or unusual lesions that may require further investigation. Anomaly detection algorithms, such as One-Class Support Vector Machines (SVM) and Isolation Forest, can help in detecting outliers or abnormal skin lesions that may potentially be malignant or indicative of a rare condition.
4. **Generative Models:** Generative models are unsupervised learning algorithms that aim to capture the underlying data distribution and generate new samples. These models, such as Variational Autoencoders (VAEs) and Generative Adversarial Networks (GANs), have shown potential in generating synthetic skin lesion images. They can be utilized to augment the limited training data, improve generalization, and enhance the performance of skin cancer classification models.

It's important to note that while unsupervised learning algorithms have valuable applications in skin cancer detection, they are often used in combination with supervised learning methods. Unsupervised techniques can aid in exploratory analysis, data preprocessing, and feature extraction, which can subsequently be utilized in supervised models for classification and prediction tasks.

*Supervised Learning.* Analysis and detection of lesions using supervised learning techniques based on ML approaches are CNN, ANN, Linear Regression (LR), RNN, Support vector regression (SVR), and Decision Tree (DT). It implements the binary classifications of images that change the appearance of the skin as benign seborrheic keratosis versus keratinocyte-carcinomas and malignant melanomas. It utilizes the ensemble based multiresolution of CNN composed of SENet, EfficientsNets, and ResNeXt for tumor detection [25]. Similarly, the transfer learning technique of AlexNet could be used to analyze images of skin lesions and melanoma, nevus, or seborrheic keratosis are classified. In classifying skin lesion images, three popular datasets of ISIC, MED-NODE, and Derm-Quest [30]. In the segmentation of affected skin regions by applying the segmentation algorithm of FRCN, CDNN was used. It uses the publicly available datasets of ISBI 2017 and PH2 [4]. Deep learning-based dilated convolutional-based transfer learning techniques of VGG16, VGG19, MobileNet, Inception-V3, and K-Means Clustering were applied for the skin cancer classification. For that, the dataset of HAM10000 with 10015 dermoscopic images [49].

Supervised learning algorithms are widely employed in skin cancer detection for their ability to learn from labeled data and make predictions based on learned patterns. These algorithms require a labeled dataset consisting of skin images with corresponding class labels (e.g., malignant or benign) to train a model. Here are some commonly used supervised learning algorithms in skin cancer detection:

1. **Support Vector Machines (SVM):** SVM is a popular algorithm that separates data points by constructing a hyperplane in a high-dimensional feature space. In skin cancer detection, SVM can be used to classify skin lesions based on extracted features. It aims to find the optimal decision boundary that maximally separates malignant and benign lesions. SVM can handle both linear and non-linear classification tasks through the use of appropriate kernel functions.
2. **Decision Trees:** Decision tree algorithms partition the data based on different features and create a hierarchical structure of decision rules. Each internal node represents a decision based on a specific feature, while leaf nodes correspond to class labels. Decision trees can be employed in skin cancer detection by considering various visual features of skin lesions, such as color, texture, and shape. They provide interpretable models that can aid dermatologists in understanding the decision-making process.
3. **Random Forests:** Random forests combine multiple decision trees to form an ensemble model. Each tree is trained on a subset of the data, and the final prediction is determined by aggregating the predictions of individual trees. Random forests are known for their ability to handle high-dimensional data and mitigate overfitting. In skin cancer detection, random forests can leverage a diverse set of features to make accurate predictions.
4. **Neural Networks:** Neural networks, particularly convolutional neural networks (CNNs), have gained significant attention in skin cancer detection due to their ability to automatically learn hierarchical representations from images. CNNs consist of multiple layers of interconnected neurons that can

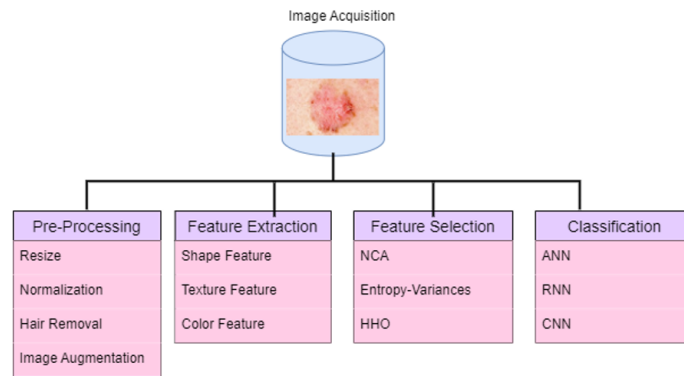


Fig. 6.1: Analysis of Image Processing Tools in Skin Cancer

extract features at different levels of abstraction. They are trained on large labeled datasets and can capture intricate patterns in skin lesions. CNNs have achieved state-of-the-art performance in skin cancer classification tasks.

5. Ensemble Methods: Ensemble methods combine multiple models to make collective predictions. Bagging and boosting are common techniques used in ensemble learning. Bagging methods, such as the Bootstrap Aggregating (or Bagging) algorithm, create multiple models trained on different subsets of the data. Boosting methods, like AdaBoost, iteratively train models, giving more weight to previously misclassified instances. Ensemble methods can enhance the accuracy and robustness of skin cancer detection models.

**6. Image Processing Techniques Skin tumor detection.** This section presents the image processing techniques required for skin lesion image analysis. The unusual growth of cells or tissue in any place of the human being is considered malignant cells or tumor cells. Digital image processing techniques are instrumental in analyzing biomedical images because they can help filter important features from the images and are used to make accurate diagnoses and treatment decisions. In the analysis of cancer images of skin with various parameters of Asymmetry, Border, Colour, and Diameter (ABCD) using, image processing tools like extraction of texture feature, color feature, and shape feature is used to detect the cancerous or non-cancerous image. Figure 6.1 analysis the image processing tools in the skin tumors classification.

The observation of Figure 6.1 shows that image processing tools like a collection of data, pre-processing, extracting features of the skin image, feature selection, and classification of images as cancerous or non-cancerous.

*Image Acquisition.* Image acquisition is collected from various sources and DL; ML algorithms are used to feature extraction of the image. Skin image datasets are collected from the internet, clinical-based images, and open dermatology databases. Expert dermatologists analyze these datasets for the removal of blurry images. Fine-tuning of skin images is collected in the dataset for further processing by applying processing tools like pre-processing, extracting features, selecting image pixels, and classification of cancer [28, 72, 23].

*Pre-Processing.* To enhance image quality, pre-processing needs to be done. It minimizes the irrelevant data from the pictures, improvising relevant data intensities and reliability. It consists of four stages: resizing, normalization, hair removal, and data augmentation.

*Resize of Image.* To accurately detect skin cancer and minimize the execution time, resizing the image is required. Skin images are resized by applying scaling and clipping procedures [39]. It resizes the images into  $224 \times 224$  or  $128 \times 128$  pixels.

*Normalization of Image .* The skin image data are normalized into the exact dimension of the interval between  $[0, 1]$  or  $[-1, 1]$ . Normalizing data is one type of linear transformation that improves skin data performance.

*Augmentation of Image.* The technique used in the augmentation process implements the methods of multiplication and transformation. In this augmentation process of the image, conversion is carried out by

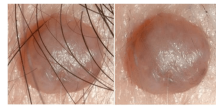


Fig. 6.2: Removal of Hair using DullRazor technique [2]

implementing a Gaussian filter. It produces the smoothing of the image and applies its displacement operators: rotation, flipping operation horizontally, zooming effect in the range of 0.2, and shearing operation in the range of 0.2 [3].

*Removal of Hair.* In diagnosing cancer in skin image, accurately and effectively remove the hair available in the skin lesion. Processing the image with hair will create confusion in extracting the feature or segmenting the affected area. Therefore, applying the DullRazor technique to remove the hair [2]. The concept behind this technique is applying a morphological operator of close to determine the hair's position in the image. Applying the substitution and bilinear interpolation operation to identify the structure of hair like long or short hair as in pixel values. Finally, applying the median filter it smoothing the image. Figure 6.2 shows the removal of hair in the skin image by applying the DullRazor technique.

*Feature Extraction.* Early diagnosis of cancer from the skin lesions are done by extracting image features. It reduces the image dimensionality which converts the pixel values into a vector format. To investigate skin cancer images with various parameters of Asymmetry, Border, Colour, and Diameter, (ABCD) using image processing tools like extraction of texture feature, color feature, and shape features are used to detect the cancerous or non-cancerous images [45]. Extracting the features of texture, color, and shape from the skin lesions using a pixel-based method. To filter out the shape features of the lesions is done by implementing the Geometric Feature. The operations in the geometric feature are area, perimeter, diameter, Irregularity Index, and Circularity Index. Texture features are extracted in skin images using a gray levels co-occurrence matrix (GLCM), LBP, and Gabor Filter. Shape feature extractions extracted from skin lesions are Wavelet transform, Fourier descriptors, Shape signature harmonic embedding, and so on. Color feature extractions are implemented by using Discrete Cosine Transform (DCT) [6].

*Feature Selection.* Feature selection is made by relevant features subset extraction of a skin image. It reduces the irrelevant features of the picture. The feature selection techniques are filter-based technique, wrapper-based, and hybrid-based process. The feature selection process is carried out by Harris Hawks optimization (HHO), Entropy-Variations, and Neighborhood Component Analysis (NCA) [31].

*Skin Cancer Classifier.* In the classification of skin cancer from large datasets by evaluating the classification algorithms of ANN, RNN, CNN with AlexNet, VGGNet, and ResNet-50. These traditional techniques are used to diagnose skin diseases accurately [38]. Computer vision technologies require an intelligent analysis model for classifying the image feature in an accurate way.

Feature selection plays a crucial role in skin cancer detection by identifying the most relevant and informative features from the input data. It aims to reduce the dimensionality of the feature space, eliminate irrelevant or redundant features, and improve the performance of the classification model. Here are some commonly used feature selection techniques in skin cancer detection:

1. Filter Methods: Filter methods evaluate the relevance of features based on statistical measures or heuristics. These methods assess each feature independently of the classification model. Common filter methods used in skin cancer detection include:
  - (a) Pearson correlation coefficient: Measures the linear correlation between features and the target class.
  - (b) Mutual information: Quantifies the mutual dependence between features and the target class.
  - (c) Chi-square test: Assesses the independence between features and the target class for categorical data.
2. Wrapper Methods: Wrapper methods evaluate subsets of features by training and testing a classification model with different combinations of features. These methods consider the interaction between features and their impact on the model's performance. Popular wrapper methods include:

Table 7.1: Summary of datasets

Dataset	Number of Images
Dermnet	23,000
DermQuest	22,082
AtlasDerm [68]	1024
PH2 [37]	200
HAM10000	10,015
ISIC archive	25,331

- (a) Recursive Feature Elimination (RFE): Starts with all features and recursively eliminates the least important ones based on model performance.
- (b) Genetic Algorithms (GA): Uses an evolutionary approach to search for an optimal feature subset by evaluating different combinations of features.
3. Embedded Methods: Embedded methods incorporate feature selection within the training process of the classification model. These methods select the most relevant features while simultaneously training the model. Some common embedded methods used in skin cancer detection include:
  - (a) L1-based regularization (e.g., Lasso or L1-regularized logistic regression): Encourages sparsity in the feature weights, effectively selecting the most informative features.
  - (b) Decision tree-based feature importance: Decision tree algorithms provide a measure of feature importance based on how much they contribute to the decision-making process.
4. Principal Component Analysis (PCA): PCA is a dimensionality reduction technique that transforms the original features into a new set of orthogonal variables, called principal components. These components capture the maximum variance in the data. In skin cancer detection, PCA can be used to reduce the dimensionality while preserving the most relevant information

**7. Data sets.** In the diagnosis of skin cancer, various algorithmic-based techniques are implemented. Evaluating and validating the detection of skin cancer results which requires the datasets collected from multiple open datasets and clinically based skin images. Training the algorithmic-based model for the classification of skin cancer by using datasets. Table 7.1 summarizes the datasets implemented in the type of skin cancer.

*Dermnet.* The Dermnet website provides information on dermatology and skin health. Dr. Thomas Habib New implemented this data source. Hampshire in the year 1998. It contains 24,000 000 or more dermoscopic pictures. And also has 643 various categories of skin diseases. These diseases are classified as two layers in the dataset. The upper layer contains 23 various types of skin diseases like melanoma, benign tumors, moles, tissue disease, and so on. The lower layer has more than 600 skin diseases.

*DermQuest.* This dataset contains 22,082 dermoscopic images. In this dataset, lesion tags are assigned to every skin lesion. Therefore, 136 lesion tags are required in the dataset.

*AtlasDerm.* This dataset is an atlas of Dermoscopic images, and it is called AtlasDerm. The images in the dataset assist dermatologists in analyzing skin cancer and detecting melanoma. It comprises 5 images of AK, BCC images 42, benign keratosis 70 images, dermatofibroma images 20, 583 melanoma, 276 melanocytic nevus, and 32 vascular skin lesions.

*PH2 Dataset.* PH2 dataset contains dermoscopic images gathered from the Center of Dermatology in Hospital Pedro Hispano, Portugal. These dataset images are 8-bit RGB color images with 768 56and pixel resolution. And also contains 200 dermoscopic images, with 40 melanoma skin cancers images, atypical nevi images 80, and 80 common nevi images. PH2 dataset contains lesion images with annotations like a clinical diagnosis of skin diseases and segmentation of medical images of pigmented skin lesions.

*HAM10000.* The publicly available dataset of skin lesions contains 11,000 training images. It includes 10,022 dermoscopic piictures gathered from the Austria-based university in Vienna called Medical University and Cliff Rosendahl's skin tumor practice at Australia Queensland. Here data is compiled by using photographic prints of lesions and digital cameras and stored in the Dermatology Department of the Medical University of

Table 7.2: ISIC Archive [16]

Dataset	Training Dataset	Testing Dataset	Total Images
ISIC_ 2016	Timng_900	Testing_ 379	Totimg_ 1279
ISIC_ 2017	Timing2000	Testing_ 600	Totimg_ 2600
ISIC_ 2018	Timing10,015	Testing_ 1512	Totimg_ 11,527
ISIC_ 2019	Timing25,331	Testing_ 8238	Totimg_ 33,569
ISIC_ 2020	Timing33,126	Testing_ 10,982	Totimg_ 44,108

Vienna, Austria. Digitized lesions photographic prints uses a Nikon-Coolsan-5000-ED scanner and convert the image into JPEG format of 8-bit images with 300 DPI quality. Cropping the image and saving it in the resolution of pixels at 72 DPI [61].

*ISIC Archive.* Data is collected from various skin lesions and stored at the ISIC archive. In 2016, the International Collaboration of Skin Imaging released ISIC dataset at the International Biomedical Imaging (ISBI) Symposium on Challenge 2016, termed ISIC 2016. The dataset of ISIC 2016 is categorized into two sections training and testing datasets. The dataset of ISIC contains 900.00 training dataset images and 379 dermoscopic image testing datasets. This data set includes two variations of skin lesion images, namely malignant melanomas and benign nevi. In the archive of ISIC, every year, many skin lesion images are included; also, 30.3% are melanoma lesions, and 69.7% are benign nevi images. These images are established for the analysis of images on skin lesions.

The archive of ISIC 2017 contains three types of images, namely benign nevi, melanomas, and seborrheic keratoses (SK). This archive includes training images of 2000, 600 testings, and validation images of 150. In this training dataset, melanoma images are 374, SK images are 254, and benign nevi images are 1372. Similarly, the dataset for validating the skin cancer images like melanoma 30 images, SK 42 images, and benign nevus 78 images. The testing dataset contains 117 melanoma pictures, SK 90 images, and benign nevus images 393.

ISIC2018 archive dataset contains a training dataset of 12,594 skin lesions, a testing dataset of 1000 images, and validation images of 100. ISIC2019 archive dataset contains 25,331 images with eight types of skin lesion images like AK, BCC, melanoma, benign keratosis, melanocytic-nevus, SCC, vascular lesion, and dermatofibroma. For the dataset for testing the 8239 images. The summary of the ISIC archive dataset is given in Table 7.2. The table shows the summary of the ISIC archive from the years 2016 to 2020. Every year, more images are included in the archive. In the ISIC 2016 archive, training images are 900, testing images are 379, and its total images are 1279. It is also classified into two varieties of malignant or benign skin lesions [26]. In the ISIC 2017 archive, training images are 2000, testing images are 600, and its total images are 2600. Skin cancers are differentiated into different categories based on the appearance of the lesion or growth. Nevus, also known as a mole, is a common type of benign skin growth (non-cancerous). Seborrheic keratosis is a soft growth that appears as a waxy or scaly patch on the skin. Melanoma is a skin cancer type characterized by the uncontrolled growth of pigment-producing cells (melanocytes) in the skin [17]. In the ISIC 2018 archive, training images are 10,015, testing images are 1512, and its total images are 11,527. It also contains melanocytic nevus, melanoma, vascular lesion, dermato fibroma, , and benign and actinic keratosis. In the ISIC 2019 archive, training images are 25,331, testing images are 8238, and its total images are 33,569. It also contains melanocytic nevus, melanoma, basal cell cancer, dermato fibroma, vascular lesions, benign keratosis, and actinic keratosis of the training dataset. The testing dataset contains 9 group classes dataset. it also includes the metadata of patient age, gender, and lesion ID [19]. In the ISIC 2020 archive, training images are 33,126; testing images are 10,982 and total images are 44,108. It is mainly focused on melanoma skin lesions. The metadata of ISIC 2020 contains the ID of the patient, ID of the lesion, gender, age, anatomical site, and diagnosis [52].

**8. Challenges in the Skin cancer examine.** This section presents those open research challenges in the diagnosis of skin lesions. It includes the factors of training the algorithmic model, variations in the size of skin lesions, light-skinned images in the dataset, subgroup class of skin lesions, an unbalanced dataset of skin

cancer images, lack of hardware, lack of age-wise categorization of skin images.

*Training the Algorithmic Model.* The significant challenges in training the algorithmic model for skin cancer identification images from the standard dataset. A minimum computation time is required to identify cancer images from dermoscopic pictures [21] effectively.

*Variations in Size of Skin Lesions.* Another research challenge in identifying and analyzing skin cancer is the dataset's image size variations. Since data is collected from various sources and varies in size. Therefore, handling the multiple dimensions of images takes a lot of work. If the image is small, it's too difficult to diagnose the earlier stage of skin cancer.

*Light Color Skinned Images in the Dataset.* The benchmark dataset contains light color -skinned patients from the United States, Australia, and Europe. The accurate and efficient detection of cancer disease of light color skinned patients required a standard algorithmic-based model [12].

*Sub-Group Class of Skin Lesions.* In the categorization of subgroups, class skin lesions are in various forms. It defines the variation between cancerous and non-cancerous skin cancer images. These minor variations in the subgroup class are difficult to predict for melanoma cancer and the birthmark of the patient [70].

*Unbalanced Dataset of Skin Cancer Image.* The collected images could be more balanced from the several sources of skin lesions. Since it contains various types of skin cancer images. It is also a great challenge to handle this unbalanced data. Unbalanced data in the dataset is difficult to generalize from the subset relevant intensity values of dermoscopic images [44].

*Detection of Skin Cancer and Hardware Issues.* Training a dataset of dermoscopic images using algorithmic-based techniques requires powerful hardware and software resources. To detect skin cancer effectively, extract unique, relevant subset pixel values from the image of the skin lesion. To do this requires powerful resources of hardware and software. This lack of hardware availability is a great challenge.

*Lack of Age-Wise Categorization of Skin Images.* The collected data of skin cancer types like BCC, SCC, and Merkel cell cancer are most probably available below age 65. The existing standard dataset contains only young-age patient skin cancer images. To accurately evaluate the process of skin cancer analysis for elderly people, images are lacking in the dataset [24].

**9. Performance of Metric Measures in Skin Cancer Images.** For the practical analysis of algorithmic-based detection of cancer in the skin by evaluating the performance of error metric measures are reviewed by Root Mean Squared Error (RMSE), Peak Signal-to-Noise Ratio (PSNR), Mean Squared Error (MSE), Structural Similarities Index (SSIM), accuracy, sensitivity, specificity.

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

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_{pi} - y_{ai})^2 \quad (9.1)$$

*RMSE.* It is similar to MSE but computes this by the root of MSE.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_{pi} - \hat{y}_{ai})^2} \quad (9.2)$$

*Peak Signal-to-Noise Ratio (PSNR).*

$$PSNR = 20 \log_{10} \left( \frac{MAX_f}{\sqrt{MSE}} \right) \quad (9.3)$$

*Structural Similarity Index (SSIM).* SSIM Index quality assessment index is a combination of three factors like luminance ( $l$ ), contrast ( $c$ ) and structure ( $s$ ) of the image.

$$SSIM(x, y) = [l(x, y)]^\alpha \times [c(x, y)]^\beta \times [s(x, y)]^\gamma \quad (9.4)$$

*Sensitivity.*

$$Sensitivity = \frac{TP}{TP + FN} \times 100 \quad (9.5)$$

*Specificity.* It is used to evaluate the rate between True Negative (TN) and True Positive (TP)

$$Specificity = \frac{TN}{TN + FP} \times 100 \quad (9.6)$$

*Accuracy.*

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100 \quad (9.7)$$

## 10. Limitations.

*Limited and Imbalanced Datasets.* Skin cancer detection often relies on labeled datasets for training machine learning models. However, obtaining large and diverse datasets with balanced representation of different skin cancer types can be challenging. Limited data can lead to overfitting, reduced generalization, and biased results. Imbalanced datasets, where one class dominates over others, can lead to poor performance in detecting minority classes.

*Variability in Image Quality.* Skin lesion images used in detection can vary in terms of image quality, lighting conditions, resolution, and image artifacts. These variations can affect the performance of the detection algorithms, making it difficult to generalize the model across different imaging devices or settings.

*Interclass Variability and Similarity.* Skin cancer encompasses various subtypes, each with unique visual characteristics. However, there can be considerable overlap and similarity between different classes, making accurate classification challenging. Distinguishing between benign and malignant lesions or differentiating between subtypes can be particularly difficult, even for expert dermatologists.

*Subjectivity and Variability in Dermatologist Diagnosis.* Dermatologists' diagnosis of skin cancer can have a level of subjectivity and inter-observer variability. This subjectivity can impact the accuracy of labeled data used for training machine learning models. Disagreements among dermatologists can lead to inconsistent labeling and ambiguous ground truth, affecting the performance of the detection algorithms.

*Limited Interpretability of Deep Learning Models.* Deep learning models, such as convolutional neural networks (CNNs), have shown impressive performance in skin cancer detection. However, their complex architectures often lack interpretability, making it challenging to understand the decision-making process or extract meaningful insights from the model's predictions. This limited interpretability can hinder the adoption and trust in deep learning-based detection systems.

*Generalization to Diverse Populations.* Skin cancer detection algorithms are typically developed and evaluated on specific populations or datasets. The performance of these algorithms on diverse populations, including different ethnicities, skin types, and age groups, may vary. Models trained on one population may not generalize well to others, leading to potential biases and reduced accuracy in real-world scenarios.

*Ethical and Privacy Concerns.* Skin cancer detection systems often require the collection and storage of sensitive medical data, including images and patient information. Ensuring the privacy and security of this data, as well as obtaining proper informed consent, is crucial. Additionally, biases in the data or algorithmic predictions could result in unfair treatment or disparities in healthcare access and outcomes.

**11. Conclusion.** This paper mainly focused on a detailed, comprehensive review of the algorithmic-based analysis of skin cancer. Initially, we surveyed related works of skin cancer detection, then discussed overview layers of skin and its diseases. This paper analyzed skin cancer identification methods and analysis of various deep learning and machine learning techniques for the study of skin cancer, in addition to this present image processing techniques for the analysis of skin cancer diseases. Another paper is focused on various datasets in cancer analysis. This survey was effectively used in skin cancer identification based on performance of metric measures of Roots Mean Squared Error (RMSE), Peak Signal-to-Noise Ratios (PSNR), Mean Squared Error (MSE), Structural Similarity (SSIM) Index, sensitivity, accuracy and specificity. These metrics are used for significantly improving the presentation of the analysis of skin cancer systems.

## REFERENCES

- [1] I. S. A. ABDELHALIM, M. F. MOHAMED, AND Y. B. MAHDY, *Data augmentation for skin lesion using self-attention based progressive generative adversarial network*, Expert Systems with Applications, 165 (2021), p. 113922.
- [2] I. ABUNADI AND E. M. SENAN, *Deep learning and machine learning techniques of diagnosis dermoscopy images for early detection of skin diseases*, Electronics, 10 (2021), p. 3158.
- [3] A. A. ADEGUN, S. VIRIRI, AND M. H. YOUSAF, *A probabilistic-based deep learning model for skin lesion segmentation*, Applied Sciences, 11 (2021), p. 3025.
- [4] M. A. AL-MASNI, M. A. AL-ANTARI, M.-T. CHOI, S.-M. HAN, AND T.-S. KIM, *Skin lesion segmentation in dermoscopy images via deep full resolution convolutional networks*, Computer methods and programs in biomedicine, 162 (2018), pp. 221–231.
- [5] M. A. ALBAHAR, *Skin lesion classification using convolutional neural network with novel regularizer*, IEEE Access, 7 (2019), pp. 38306–38313.
- [6] I. ALFI, A.; rahman, mm; shorfuzzaman, m.; nazir, A. A Non-Invasive Interpretable Diagnosis of Melanoma Skin Cancer Using Deep Learning and Ensemble Stacking of Machine Learning Models. Diagnostics, 12 (2022), p. 726.
- [7] A.-R. ALI, J. LI, G. YANG, AND S. J. O'SHEA, *A machine learning approach to automatic detection of irregularity in skin lesion border using dermoscopic images*, PeerJ Computer Science, 6 (2020), p. e268.
- [8] R. ASHRAF, S. AFZAL, A. U. REHMAN, S. GUL, J. BABER, M. BAKHTYAR, I. MEHMOOD, O.-Y. SONG, AND M. MAQSOOD, *Region-of-interest based transfer learning assisted framework for skin cancer detection*, IEEE Access, 8 (2020), pp. 147858–147871.
- [9] M. ATTIA, M. HOSSNY, H. ZHOU, S. NAHAVANDI, H. ASADI, AND A. YAZDABADI, *Realistic hair simulator for skin lesion images: A novel benchmarking tool*, Artificial Intelligence in Medicine, 108 (2020), p. 101933.
- [10] L. BI, J. KIM, E. AHN, A. KUMAR, D. FENG, AND M. FULHAM, *Step-wise integration of deep class-specific learning for dermoscopic image segmentation*, Pattern recognition, 85 (2019), pp. 78–89.
- [11] D. BISLA, A. CHOROMANSKA, R. S. BERMAN, J. A. STEIN, AND D. POLSKY, *Towards automated melanoma detection with deep learning: Data purification and augmentation*, in Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops, 2019, pp. 0–0.
- [12] T. J. BRINKER, A. HEKLER, J. S. UTIKAL, N. GRABE, D. SCHADENDORF, J. KLODE, C. BERKING, T. STEEB, A. H. ENK, AND C. VON KALLE, *Skin cancer classification using convolutional neural networks: systematic review*, Journal of medical Internet research, 20 (2018), p. e11936.
- [13] G. A. R. BRUNN, D. N. BRANDÃO, R. S. RIBEIRO, A. M. P. F. CARVALHO, M. Â. BRANDRÃO, ET AL., *Primary breast parenchyma melanoma: case report*, Brazilian Journal of Oncology, 17 (2021), pp. 1–3.
- [14] P. BUMRUNKUN, K. CHAMNONGTHAI, AND W. PATCHOO, *Initial geometrical templates with parameter sets for active contour on skin cancer boundary segmentation*, Journal of Healthcare Engineering, 2021 (2021).
- [15] A. L. BYRD, Y. BELKAID, AND J. A. SEGRE, *The human skin microbiome*, Nature Reviews Microbiology, 16 (2018), pp. 143–155.
- [16] B. CASSIDY, C. KENDRICK, A. BRODZICKI, J. JAWOREK-KORJAKOWSKA, AND M. H. YAP, *Analysis of the isic image datasets: Usage, benchmarks and recommendations*, Medical image analysis, 75 (2022), p. 102305.
- [17] N. C. CODELLA, D. GUTMAN, M. E. CELEBI, B. HELBA, M. A. MARCHETTI, S. W. DUSZA, A. KALLOO, K. LIOPYRIS, N. MISHRA, H. KITTLER, ET AL., *Skin lesion analysis toward melanoma detection: A challenge at the 2017 international symposium on biomedical imaging (isbi), hosted by the international skin imaging collaboration (isic)*, in 2018 IEEE 15th international symposium on biomedical imaging (ISBI, IEEE, 2018, pp. 168–172.
- [18] N. C. CODELLA, Q.-B. NGUYEN, S. PANKANTI, D. A. GUTMAN, B. HELBA, A. C. HALPERN, AND J. R. SMITH, *Deep learning ensembles for melanoma recognition in dermoscopy images*, IBM Journal of Research and Development, 61 (2017), pp. 5–1.
- [19] M. COMBALIA, N. C. CODELLA, V. ROTEMBERG, B. HELBA, V. VILAPLANA, O. REITER, C. CARRERA, A. BARREIRO, A. C. HALPERN, S. PUIG, ET AL., *Bcn20000: Dermoscopic lesions in the wild*, arXiv preprint arXiv:1908.02288, (2019).
- [20] K. DAS, C. J. COCKERELL, A. PATIL, P. PIETKIEWICZ, M. GIULINI, S. GRABBE, AND M. GOLDUST, *Machine learning and its application in skin cancer*, International Journal of Environmental Research and Public Health, 18 (2021), p. 13409.
- [21] M. DILDAR, S. AKRAM, M. IRFAN, H. U. KHAN, M. RAMZAN, A. R. MAHMOOD, S. A. ALSAIARI, A. H. M. SAEED, M. O. ALRADDADI, AND M. H. MAHNASHI, *Skin cancer detection: a review using deep learning techniques*, International journal of environmental research and public health, 18 (2021), p. 5479.
- [22] U.-O. DORJ, K.-K. LEE, J.-Y. CHOI, AND M. LEE, *The skin cancer classification using deep convolutional neural network*, Multimedia Tools and Applications, 77 (2018), pp. 9909–9924.
- [23] Y. FUJISAWA, Y. OTOMO, Y. OGATA, Y. NAKAMURA, R. FUJITA, Y. ISHITSUKA, R. WATANABE, N. OKIYAMA, K. OHARA, AND M. FUJIMOTO, *Deep-learning-based, computer-aided classifier developed with a small dataset of clinical images surpasses board-certified dermatologists in skin tumour diagnosis*, British Journal of Dermatology, 180 (2019), pp. 373–381.
- [24] S. GARCOVICH, G. COLLOCA, P. SOLLENA, B. ANDREA, L. BALDUCCI, W. C. CHO, R. BERNABEI, AND K. PERIS, *Skin cancer epidemics in the elderly as an emerging issue in geriatric oncology*, Aging and disease, 8 (2017), p. 643.
- [25] N. GESSERT, M. NIELSEN, M. SHAIKH, R. WERNER, AND A. SCHLAEFER, *Skin lesion classification using ensembles of multi-resolution efficientnets with meta data*, MethodsX, 7 (2020), p. 100864.
- [26] D. GUTMAN, N. C. CODELLA, E. CELEBI, B. HELBA, M. MARCHETTI, N. MISHRA, AND A. HALPERN, *Skin lesion analysis toward melanoma detection: A challenge at the international symposium on biomedical imaging (isbi) 2016, hosted by the international skin imaging collaboration (isic)*, arXiv preprint arXiv:1605.01397, (2016).
- [27] H. A. HAENSSLE, C. FINK, R. SCHNEIDERBAUER, F. TOBERER, T. BUHL, A. BLUM, A. KALLOO, A. B. H. HASSEN, L. THOMAS,



- A. ENK, ET AL., *Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists*, *Annals of oncology*, 29 (2018), pp. 1836–1842.
- [28] S. S. HAN, M. S. KIM, W. LIM, G. H. PARK, I. PARK, AND S. E. CHANG, *Classification of the clinical images for benign and malignant cutaneous tumors using a deep learning algorithm*, *Journal of Investigative Dermatology*, 138 (2018), pp. 1529–1538.
- [29] A. HEKLER, J. S. UTIKAL, A. H. ENK, A. HAUSCHILD, M. WEICHTHAL, R. C. MARON, C. BERKING, S. HAFERKAMP, J. KLODE, D. SCHADENDORF, ET AL., *Superior skin cancer classification by the combination of human and artificial intelligence*, *European Journal of Cancer*, 120 (2019), pp. 114–121.
- [30] K. M. HOSNY, M. A. KASSEM, AND M. M. FOAUD, *Classification of skin lesions using transfer learning and augmentation with alex-net*, *PloS one*, 14 (2019), p. e0217293.
- [31] A. A. JIMÉNEZ, F. P. G. MÁRQUEZ, V. B. MORALEDA, AND C. Q. G. MUÑOZ, *Linear and nonlinear features and machine learning for wind turbine blade ice detection and diagnosis*, *Renewable energy*, 132 (2019), pp. 1034–1048.
- [32] M. F. JOJOA ACOSTA, L. Y. CABALLERO TOVAR, M. B. GARCIA-ZAPIRAIN, AND W. S. PERCYBROOKS, *Melanoma diagnosis using deep learning techniques on dermoscopic images*, *BMC Medical Imaging*, 21 (2021), pp. 1–11.
- [33] P. KAVITHA AND V. JAYALAKSHMI, *Survey of skin cancer detection using various image processing techniques*, in 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS), IEEE, 2020, pp. 1062–1069.
- [34] M. Q. KHAN, A. HUSSAIN, S. U. REHMAN, U. KHAN, M. MAQSOOD, K. MEHMOOD, AND M. A. KHAN, *Classification of melanoma and nevus in digital images for diagnosis of skin cancer*, *IEEE Access*, 7 (2019), pp. 90132–90144.
- [35] D. KIM AND B.-W. HONG, *Unsupervised feature elimination via generative adversarial networks: application to hair removal in melanoma classification*, *IEEE Access*, 9 (2021), pp. 42610–42620.
- [36] A. R. LAKSHMINARAYANAN, R. BHUVANESHWARI, S. BHUVANESHWARI, S. PARTHASARATHY, S. JEGANATHAN, AND K. M. SAGAYAM, *Skin cancer prediction using machine learning algorithms*, in *Artificial Intelligence and Technologies: Select Proceedings of ICRTAC-AIT 2020*, Springer, 2022, pp. 303–310.
- [37] H. LIAO, Y. LI, AND J. LUO, *Skin disease classification versus skin lesion characterization: Achieving robust diagnosis using multi-label deep neural networks*, in 2016 23rd International Conference on Pattern Recognition (ICPR), IEEE, 2016, pp. 355–360.
- [38] I. E. LIVIERIS, *Improving the classification efficiency of an ann utilizing a new training methodology*, in *Informatics*, vol. 6, MDPI, 2018, p. 1.
- [39] A. MAHBOD, G. SCHAEFER, C. WANG, G. DORFFNER, R. ECKER, AND I. ELLINGER, *Transfer learning using a multi-scale and multi-network ensemble for skin lesion classification*, *Computer methods and programs in biomedicine*, 193 (2020), p. 105475.
- [40] A. MAHBOD, G. SCHAEFER, C. WANG, R. ECKER, AND I. ELLINGE, *Skin lesion classification using hybrid deep neural networks*, in *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2019, pp. 1229–1233.
- [41] S. MAHESWARAN, S. SATHESH, P. PRIYADHARSHINI, AND B. VIVEK, *Identification of artificially ripened fruits using smart phones*, in 2017 international conference on intelligent computing and control (I2C2), IEEE, 2017, pp. 1–6.
- [42] S. MAHESWARAN, S. SATHESH, G. SARAN, AND B. VIVEK, *Automated coconut tree climber*, in 2017 International Conference on Intelligent Computing and Control (I2C2), IEEE, 2017, pp. 1–6.
- [43] D. B. MENDES AND N. C. DA SILVA, *Skin lesions classification using convolutional neural networks in clinical images*, *arXiv preprint arXiv:1812.02316*, (2018).
- [44] M. A. A. MILTON, *Automated skin lesion classification using ensemble of deep neural networks in isic 2018: Skin lesion analysis towards melanoma detection challenge*, *arXiv preprint arXiv:1901.10802*, (2019).
- [45] M. NASIR, M. ATTIQUE KHAN, M. SHARIF, I. U. LALI, T. SABA, AND T. IQBAL, *An improved strategy for skin lesion detection and classification using uniform segmentation and feature selection based approach*, *Microscopy research and technique*, 81 (2018), pp. 528–543.
- [46] N. NIDA, A. IRTAZA, A. JAVED, M. H. YOUSAF, AND M. T. MAHMOOD, *Melanoma lesion detection and segmentation using deep region based convolutional neural network and fuzzy c-means clustering*, *International journal of medical informatics*, 124 (2019), pp. 37–48.
- [47] S. PATHAN, K. G. PRABHU, AND P. SIDDALINGASWAMY, *Techniques and algorithms for computer aided diagnosis of pigmented skin lesions—a review*, *Biomedical Signal Processing and Control*, 39 (2018), pp. 237–262.
- [48] P. M. PEREIRA, L. A. THOMAZ, L. M. TAVORA, P. A. ASSUNCAO, R. FONSECA-PINTO, R. P. PAIVA, AND S. M. FARIA, *Multiple instance learning using 3d features for melanoma detection*, *IEEE Access*, 10 (2022), pp. 76296–76309.
- [49] M. A. R. RATUL, M. H. MOZAFFARI, W.-S. LEE, AND E. PARIMBELLI, *Skin lesions classification using deep learning based on dilated convolution*, *BioRxiv*, (2019), p. 860700.
- [50] K. RAZA AND N. K. SINGH, *A tour of unsupervised deep learning for medical image analysis*, *Current Medical Imaging*, 17 (2021), pp. 1059–1077.
- [51] A. REZVANTALAB, H. SAFIGHOLI, AND S. KARIMIJEHNI, *Dermatologist level dermoscopy skin cancer classification using different deep learning convolutional neural networks algorithms*, *arXiv preprint arXiv:1810.10348*, (2018).
- [52] V. ROTEMBERG, N. KURTANSKY, B. BETZ-STABLEIN, L. CAFFERY, E. CHOUSAKOS, N. CODELLA, M. COMBALIA, S. DUSZA, P. GUITERA, D. GUTMAN, ET AL., *A patient-centric dataset of images and metadata for identifying melanomas using clinical context*, *Scientific data*, 8 (2021), p. 34.
- [53] F. RUNDO, G. L. BANNA, AND S. CONOCI, *Bio-inspired deep-cnn pipeline for skin cancer early diagnosis*, *Computation*, 7 (2019), p. 44.
- [54] A. SAGAR AND J. DHEEBA, *Convolutional neural networks for classifying melanoma images*, *BioRxiv*, (2020), pp. 2020–05.
- [55] P. SAJID AND D. RAJESH, *Performance evaluation of classifiers for automatic early detection of skin cancer*, *J. Adv. Res.*

- Dyn. Control. Syst, 10 (2018), pp. 454–461.
- [56] F. N. SHAIK AND A. NAAZ, *Skin cancer detection using neural networks*.
- [57] R. L. SIEGEL, K. D. MILLER, AND A. JEMAL, *Cancer statistics, 2018*, CA: a cancer journal for clinicians, 68 (2018), pp. 7–30.
- [58] T. SKUHALA, V. TRKULJA, M. RIMAC, A. DRAGOBRATOVIĆ, AND B. DESNICA, *Analysis of types of skin lesions and diseases in everyday infectious disease practice—how experienced are we?*, Life, 12 (2022), p. 978.
- [59] K. SRINIDHI, M. ASHWINI, B. S. SREE, P. APOORVA, V. AKSHITHA, AND K. KHASIM, *Brain tumor segmentation and survival prediction using 3d u-net*, in Journal of Physics: Conference Series, vol. 2325, IOP Publishing, 2022, p. 012048.
- [60] P. S. STAFF, *What is the difference between melanoma and non-melanoma skin cancer*, PSS. Available online: <https://www.premiersurgical.com/01/whats-the-difference-between-melanoma-and-non-melanoma-skin-cancer/> (accessed on 6 February).
- [61] P. TSCHANDL, C. ROSENDAHL, AND H. KITTLER, *The ham10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions*, Scientific data, 5 (2018), pp. 1–9.
- [62] M. VIJAYALAKSHMI, *Melanoma skin cancer detection using image processing and machine learning*, International Journal of Trend in Scientific Research and Development (IJTSRD), 3 (2019), pp. 780–784.
- [63] B. VIVEK, S. MAHESWARAN, P. KEERTHANA, S. SATHESH, S. BRINGERAJ, R. A. SRI, AND S. A. SULTHANA, *Low-cost raspberry pi oscilloscope*, in 2018 international conference on intelligent computing and communication for smart World (I2C2SW), IEEE, 2018, pp. 386–390.
- [64] Y. WANG, D. C. LOUIE, J. CAI, L. TCHVIALEVA, H. LUI, Z. J. WANG, AND T. K. LEE, *Deep learning enhances polarization speckle for in vivo skin cancer detection*, Optics & Laser Technology, 140 (2021), p. 107006.
- [65] L. WEI, K. DING, AND H. HU, *Automatic skin cancer detection in dermoscopy images based on ensemble lightweight deep learning network*, IEEE Access, 8 (2020), pp. 99633–99647.
- [66] C. XIA, X. DONG, H. LI, M. CAO, D. SUN, S. HE, F. YANG, X. YAN, S. ZHANG, N. LI, ET AL., *Cancer statistics in china and united states, 2022: profiles, trends, and determinants*, Chinese medical journal, 135 (2022), pp. 584–590.
- [67] M. YACIN SIKKANDAR, B. A. ALRASHEADI, N. PRAKASH, G. HEMALAKSHMI, A. MOHANARATHINAM, AND K. SHANKAR, *Deep learning based an automated skin lesion segmentation and intelligent classification model*, Journal of ambient intelligence and humanized computing, 12 (2021), pp. 3245–3255.
- [68] Y. YAN, G. SHEN, S. ZHANG, T. HUANG, Z.-H. DENG, AND U. YUN, *Sequence generative adversarial nets with a conditional discriminator*, Neurocomputing, 429 (2021), pp. 69–76.
- [69] S. YERUSHALMI, J. Y. SHIRAZI, AND A. FRIEDMAN, *New developments in bacterial, viral, and fungal cutaneous infections*, Current dermatology reports, 9 (2020), pp. 152–165.
- [70] L. YU, H. CHEN, Q. DOU, J. QIN, AND P.-A. HENG, *Automated melanoma recognition in dermoscopy images via very deep residual networks*, IEEE transactions on medical imaging, 36 (2016), pp. 994–1004.
- [71] B. ZHANG, X. ZHOU, Y. LUO, H. ZHANG, H. YANG, J. MA, AND L. MA, *Opportunities and challenges: Classification of skin disease based on deep learning*, Chinese Journal of Mechanical Engineering, 34 (2021), pp. 1–14.
- [72] J. ZHANG, Y. XIE, Y. XIA, AND C. SHEN, *Attention residual learning for skin lesion classification*, IEEE transactions on medical imaging, 38 (2019), pp. 2092–2103.
- [73] L. ZHANG, H. J. GAO, J. ZHANG, AND B. BADAMI, *Optimization of the convolutional neural networks for automatic detection of skin cancer*, Open Medicine, 15 (2020), pp. 27–37.

*Edited by:* Sathishkumar V E

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

*Received:* Apr 16, 2023

*Accepted:* Jun 24, 2023



## GENERAL LAYOUT PLANNING MODEL OF LANDSCAPE CERAMIC SCULPTURE BASED ON NSGA - II ALGORITHM

**Abstract.** The current overall layout planning model matrix of landscape ceramic sculpture is generally unidirectional, and the planning efficiency is low, resulting in a decline in the layout optimization ratio of the model. Therefore, the design and verification analysis of landscape ceramic sculpture's overall layout planning model based on the Nondominated Sorting Genetic Algorithm (NSGA - II) algorithm is proposed. According to the actual planning needs and standards, first set the basic layout points, establish a cross-planning matrix in a multi-level manner, and improve the efficiency of the overall layout planning of the sculpture. The NSGA - II calculation landscape ceramic sculpture layout planning structure is constructed on this basis, and the model design is realized by level conversion. This novel NSGA-II with level conversion performs better layout planning when compared with other conventional models. The final test results show that through three stages of layout optimization processing, compared with the initial planning layout, the optimal layout optimization ratio for the setting of the plaza sculpture can reach more than 60%, indicating that with the help of this method, the layout planning of sculpture has been further improved, the space has been expanded, and has practical application value.

**Key words:** NSGA - II algorithm; Landscape ceramics; Ceramic sculpture; General layout; Planning model; Layout planning;

**1. Introduction.** Most of the landscape ceramic sculptures are large urban sculptures, which are generally set in a more prominent urban public environment and are symbols of the city or a particular area, such as large squares, public green spaces, public buildings and other areas [13]. Landscape ceramic sculpture is usually large in scale and system, so there are many factors to consider in planning and design [22]. The outdoor landscape ceramic sculpture is also an essential part of modern cities to some extent and needs to be integrated with the town when planning [11]. The styles of landscape ceramic sculptures are different and varied. Generally, they are directly related to their location. They can be roughly divided into the following categories: memorial, thematic, decorative, entertaining, and religious [6]. There are also significant differences in the overall layout planning of the built-in design of different landscape ceramic sculptures. As a public environment art, the comprehensive layout and overall planning directly affect the design effect in the later period [17]. To strengthen the artistry of sculpture, relevant staff usually design in advance through the overall layout planning model before manufacturing landscape ceramic sculpture and deal with the details of sculpture [2]. However, most of the traditional sculpture general layout planning models are one-way. Although they can achieve the expected planning tasks and goals, they lack pertinence and stability. Sculptures are also prone to problems in the production process, such as inconsistency with the surrounding environment and planning chaos [9]. Therefore, the design and verification analysis of landscape ceramic sculpture's overall layout planning model based on the NSGA - II algorithm is proposed. The so-called NSGA - II algorithm is one of the multi-objective genetic algorithms. It has a strong pertinence in calculation. In application, it uses diversified methods to reduce the circularity and complexity of the non-inferior sorting genetic algorithm, further strengthens the actual running speed to a certain extent, and has the characteristics of good convergence, which is convenient for later calculation of solution set [16]. Integrating this algorithm with the design of the overall layout planning model of landscape ceramic sculpture can better expand the actual planning scope, strengthen the applicability of the model, gradually build a more stable and diversified layout planning model, promote the design of the model to further integrate with the urban space environment and humanistic environment, and highlight the urban characteristic culture [12]. In addition, the application of the NSGA - II algorithm can also minimize the errors in the design process of landscape ceramic sculpture, improve the overall application performance

of the planning model, facilitate the hierarchical control of urban design, form a scientific planning guidance, have important value in the visual landscape environment design, and lay a solid foundation for the innovation and maintenance of subsequent landscape ceramic sculpture [21].

This research finds the solution for following question,

1. How the problem in Automated layout planning is formalized?
2. Layout planning requires multi objective algorithm, how the algorithm achieves the efficiency?

## 2. Design the overall layout of landscape ceramic sculpture NSGA - II calculation planning model.

**2.1. Basic layout point setting.** In general, the layout planning of landscape ceramic sculpture needs to consider the surrounding environment and fusion background and try to increase the combination degree of sculpture and environment, showing a sense of hierarchy [4]. Therefore, before designing the overall layout planning model of landscape ceramic sculpture, it is necessary to set the specific space size, layout scale, and planning location of the sculpture and calibrate the basic layout point [14]. First, measure and calculate the space proportion of the sculpture and generally control it within one-tenth of the one-way control range [8]. Subsequently, the setting of planning layout points must be integrated with the urban functional structure to form an orderly spatial sequence and reasonably control the spacing of issues, as shown in Formula 2.1 below:

$$G = \frac{\pi \times \mathfrak{R}}{f - \sum_{r=1} \pi r + \chi^2} - \mathfrak{R}(\mathfrak{N} + r)^2 \quad (2.1)$$

In Formula 2.1  $G$  represents the spacing between points,  $f$  represents a conversion sequence,  $\pi$  indicates the unit control value,  $r$  represents the number of points,  $\chi$  indicates the sensing range,  $\mathfrak{R}$  indicates the remaining space. According to the above settings, the calculation of the spacing between the points is completed. Next, the spacing value obtained from the comprehensive measurement is used to adjust the position of the initial points and reasonably expand the layout space of the sculpture [19]. Then, based on this, the area and scope set by the sculpture, combined with the characteristics of the surrounding environment, set nodes, portals, and demarcated specific fields, further highlighting the landmark perception of the sculpture from multiple aspects [23].

In addition, the layout planning and point setting of landscape ceramic sculpture also need to build a highly coordinated landscape axis and green channel, which can balance the layout of the entire sculpture, increase the surrounding adjustable space, and facilitate viewing [7]. However, it should be noted that the setting of the sculpture layout points must be coordinated with their positions and kept on the same level. To some extent, it is also conducive to the transformation of space or layout and the close integration of life [1].

**2.2. Establish multi-level cross-planning matrix.** After setting the essential layout points, the next step is integrating the NSGA - II algorithm to establish a multi-level cross-planning matrix. Different from the traditional layout planning matrix, the multi-level cross matrix designed this time covers a relatively large area and is more targeted. In the practical application process, it has a certain degree of assistance for the overall layout planning of sculpture [3]. First of all, the actual setting area of the sculpture is integrated for hierarchical control, and the public space is also divided into primary and secondary. Therefore, the model matrix needs to follow this principle in the construction process, design a cross hierarchical control structure, set a core or planning key position in the model matrix, and take this as the axis to extend to both sides, form two controllable two-way balance spaces [5]. The layout is divided and controlled at a hierarchical level, and the weight coefficient of cross-planning is calculated, as shown in Formula 2.2:

$$D = \sum_{w=1} \lambda w + \xi(b - c)^2 \quad (2.2)$$

In Equation 2.2  $D$  represents the weight coefficient of cross planning,  $\lambda$  indicates the reserved one-way distance,  $w$  indicates the number of crossings,  $\xi$  represents the displacement distance of the point,  $b$  is the scale volume ratio,  $c$  indicates the deviation of the drawing. According to the above measurement, the weight coefficient of cross-planning is calculated. This value is then set in the matrix to form a circular layout planning

Table 2.1: Classification of sculpture layout levels

<i>Sculpture layout level</i>	<i>Definition</i>	<i>Matrix Planning Location</i>
Sculpture Core Area	Main setting area of sculpture within 1 meter	Sculpture base and core points
Sculpture layout corridor	Transition corridor	Corridor length and location
Sculpture landscape area	Visiting points and surrounding areas	Landscape extension area

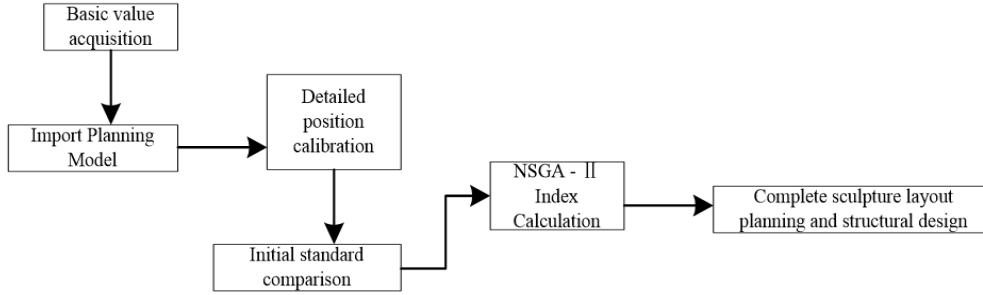


Fig. 2.1: NSGA - II Calculation Landscape Ceramic Sculpture Layout Planning Structure Diagram

structure, and the directional limit standard [15] is correspondingly configured. At this time, using the design matrix, it is necessary to divide the key control areas at the location of the sculpture layout and, at the same time, mark out the corresponding levels. The specific location is the sculpture core area, layout corridor, and sculpture landscape area [24]. The specific definition and matrix planning location are shown in Table 2.1.

According to Table 2.1, complete the division of sculpture layout levels. On this basis, adjust the planning structure of the matrix to form a more flexible and changeable layout planning framework, make directional planning for the layout of sculptures from multiple aspects, and strengthen the application ability of the matrix [17].

**2.3. Construct NSGA - II to calculate the layout planning structure of landscape ceramic sculpture.** The multi-level cross-planning matrix of the above design is imported into the layout planning structure of landscape ceramic sculpture to control and guide the sculpture construction. After completing the basic planning, the next step is to integrate the pressure setting requirements of sculpture and the collected data information [18]. Measure the scale of the single sculpture, and calculate the volume value of the sculpture using the NSGA - II algorithm, as shown in Formula 2.3:

$$Q = \int (x - z)^2 \times \omega + \frac{\sqrt{j-1}}{z\omega^2} \quad (2.3)$$

In Formula 2.3,  $Q$  represents the mass value of the sculpture,  $x$  represents the conversion ratio of the matrix,  $z$  indicates the degree ratio,  $\omega$  indicates the seating range of the sculpture,  $j$  indicates the overlap range. According to the above settings, the calculation of the sculpture mass value is completed. This value is set inside the model as the layout solidification limit standard value of the model [20]. Next, import the initial data and information of landscape ceramic sculpture into the automatic layout planning program, and design an inferior model planning structure based on this, as shown in Figure 2.1.

According to Figure 2.1, complete the design and application analysis of NSGA - II measuring landscape ceramic sculpture layout planning structure. At the same time, using the planning matrix of the design, we can re-calculate the indicators, compare the standards of the delicate areas of the sculpture layout, and obtain a scientific planning scheme to facilitate the layout planning work [10].

Table 2.2: Setting Table of Level Conversion Values

<i>Conversion indicators</i>	<i>Initial level</i>	<i>Transition level</i>
Spatial sequence	Single sequence	Bidirectional sequence
Sculpture position deviation	1.05	3.05
Local area delimitation	unify	multiple target

Here is a step-by-step approach for implementing the NSGA-II algorithm:

1. Initialize the population: Create an initial population of candidate solutions.
2. Evaluate the population: Evaluate the objective values for each individual.
3. Perform non-dominated sorting to classify the solutions into different fronts or levels based on their dominance relationships.
4. Calculate the crowding distance for each individual within each front.
5. Select parents: Select parents for the next generation using a combination of non-dominated sorting and crowding distance.
6. Reproduction: Apply genetic operators such as crossover and mutation to create offspring from the selected parents. These genetic operators introduce variation in the population and help explore different regions of the search space.
7. Create the next generation: Combine the parents and offspring to create the next generation of the population.
8. Repeat the evaluation steps 2-7, non-dominated sorting, crowding distance calculation, parent selection, reproduction, and generation creation steps for a certain number of generations or until a termination condition is met.
9. Terminate the algorithm if the condition is met.
10. Select the final Pareto front.

**2.4. Hierarchical transformation implementation model design.** The so-called hierarchical transformation mainly refers to the overall layout planning model of sculpture, it is often re-planned due to the change of sculpture points and the adjustment of hierarchical control processing standards, but this form will reduce the planning efficiency. Therefore, the hierarchical transformation method is adopted to increase the transformation ability of the model and complete the layout planning task faster. Set the hierarchy and conversion indicators, as shown in Table 2.2.

Set the value of cascade conversion according to Table 2.2. According to the planning requirements of the sculpture layout and the specific refining standards, the index parameters are converted one by one at each level, and the general situation of the sculpture setting is reasonably adjusted through the model to ensure that the optimal layout planning scheme is obtained.

**3. Experiment.** This time is mainly to analyze and verify the application effect of landscape ceramic sculpture's overall layout planning model based on the NSGA - II algorithm. Considering the authenticity and reliability of the final test results, the analysis is carried out by comparison, and the square sculpture project in urban area D is selected as the main target of the test. Use professional equipment and devices to collect primary data on landscape ceramic sculpture and information and summarize and integrate them for future use. According to the overall design requirements and standard changes of the landscape ceramic sculpture planning model, the final test results are compared and studied. Next, integrate the NSGA - II algorithm to build and correlate the test environment.

**3.1. Experiment preparation.** Based on the NSGA - II algorithm, the actual test and analysis environment of the square sculpture project in urban area D was verified and studied. City D is located in southwest China, with a total area of 9655 square kilometers. The urban environment is good, there are many squares, and the layout is complicated. It needs multi-dimensional planning. The sculpture designed this time is the theme sculpture of a square. The geographical location of the square is superior. It is located in the center of the city, with convenient transportation. It is planned to form a three-dimensional transportation network.

Table 3.1: Setting Table of Phased Layout Conditions for Sculpture in Urban Square D

<i>Sculpture design stage</i>	<i>Layout conditions</i>	<i>Proportion of controllable completion space</i>
Positioning stage	The structure matches the spatial sequence, forming a layered sense of sculpture space	3.2
Design phase	Set main nodes and key points to highlight carving Ability to perceive landmarks in sculpture works	4.1
Manufacturing phase	Closely integrate with public space and integrate with civil life	5.5

In such an environment, the selection of square sculpture needs to be atmospheric, simple, and with urban characteristics. When designing, the surrounding environment and basic settings of the square should also be considered to ensure the spatial basis of the planning of the ceramic sculpture system. According to the actual planning needs and standards, first design the design phase of the comprehensive sculpture, and then make unconditional settings, as shown in Table 3.1.

According to Table 3.1, complete the setting of phased layout conditions for sculptures in urban square D, and convert the above information into data packages or instructions into sculpture planning models for future use. In this background environment, according to the space of the square, use the sculpture planning model to calculate the floor area of the sculpture, and at the same time, mark the corresponding falling point at the setting point, so as to facilitate the setting of later sculptures. Then, on this basis, the NSGA - II algorithm is integrated to calculate the initial scale value of the sculpture, as shown in Formula 3.1:

$$H = (1 - m) \times \sum_{u=1} nu + v^2 \times \frac{1}{m + n} \quad (3.1)$$

In Formula 3.1  $H$  represents the initial scale value of the sculpture,  $m$  represents the setting space of the sculpture,  $n$  represents the controllable coverage area of the scheduling,  $u$  represents the number of landing positions of sculptures,  $v$  represents a single value conversion distance. According to the above settings, the construction and correlation of the basic test environment are completed. Next, the NSGA - II algorithm is integrated to carry out a specific measurement and analysis of the sculpture planning model of the square in D urban area.

**3.2. Experimental process and result analysis.** In the above-built test environment, the NSGA - II algorithm is integrated to carry out specific verification and analysis on the sculpture planning model of D City Square. The model designed this time is a multi-level composite landscape ceramic sculpture with strong urban cultural characteristics. First of all, the phased layout conditions and construction standards of the square sculpture set above are introduced into the model. Considering the changes and adjustments of the surrounding environment of the square, the specific floor space and area of landscape ceramic sculpture are determined first. Then, the layout planning of the sculptures will be carried out. Generally, in order to highlight the characteristics of sculptures, sculptures will be set between squares. The specific planning is shown in Figure 3.1.

According to Figure 3.1, complete the design and analysis of the initial planning structure of the sculpture in the square of urban area D. Next, synthesize the designed sculpture planning model, determine the weight proportion of the sculpture at this time, and apply the analytic hierarchy process and progressive analysis method to set the overall layout of the cloth sculpture at this time. Adjust the basic index parameters of the model at this time by constructing a judgment matrix and model, And reset some values, as shown in Table 3.2.

Set the initial index parameters of the sculpture planning model according to Table 3.2. Next, according to the above model setting conditions and sculpture manufacturing processing, complete the sculpture setting. Five local locations of the sculpture are selected for measurement. According to the initial layout standard, the overall layout planning model of the sculpture and the NSGA - II algorithm is integrated to calculate the

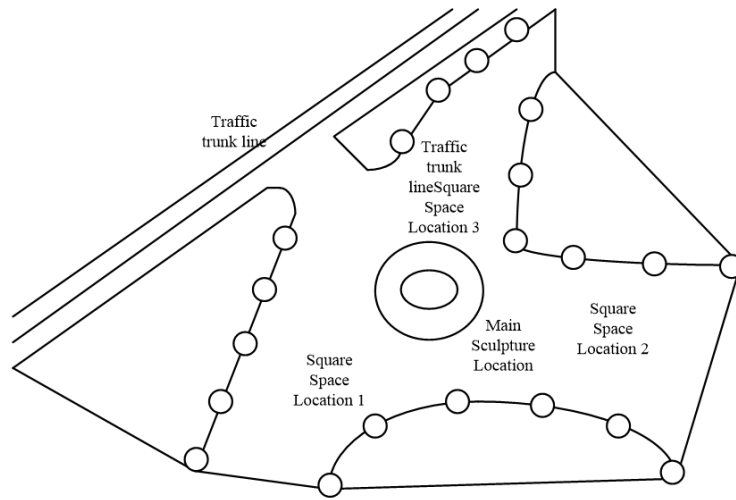


Fig. 3.1: Diagram of Initial Planning Structure of Sculpture in City Square of D

Table 3.2: Initial Index Parameter Setting of Sculpture Planning Model

<i>Initial indicators of sculpture planning model</i>	<i>Initial setting of index parameter values</i>	<i>Measured index parameter value standard</i>
Landing point setting	Set double landing points, but not stable	Set multiple landing points, with strong stability and safety
Space percentage	3.2	4.1
Preset model planning deviation	1.03	1.01
Model layout settings	18.61	16.05
Local definition of sculpture layout planning	Local layout planning is stable and balanced, with controllable spatial adjustment	The local layout must have a strong sense of hierarchy, echoing and integrating with the surrounding environment

layout optimization ratio, as shown in Formula 3.2:

$$M = O^2 \times \mathfrak{S} + (\nu + \sum_{y=1} \mathfrak{S}y - 1)^2 \tag{3.2}$$

In Formula 3.2  $M$  represents the layout optimization ratio,  $O$  represents the total area of the square,  $\mathfrak{S}$  indicates directional coverage,  $\nu$  represents the conversion distance,  $y$  indicates the measurement frequency of the model. According to the above determination, complete the analysis of the test results, as shown in Figure 3.2.

According to Figure 3.2, the comparative analysis of the test results is completed: through three stages of layout optimization processing, compared with the initial planning layout, the optimal layout optimization ratio for the setting of the square sculpture can reach more than 60%, which shows that with the help of this method, the layout planning of sculpture has been further improved, the space has been expanded, and has practical application value.

**4. Conclusion.** The research presented in this paper focuses on the design and application analysis of the overall layout planning model of landscape ceramic sculpture using the NSGA-II algorithm. The developed model demonstrates enhanced pertinence and stability compared to the original layout planning model, offering



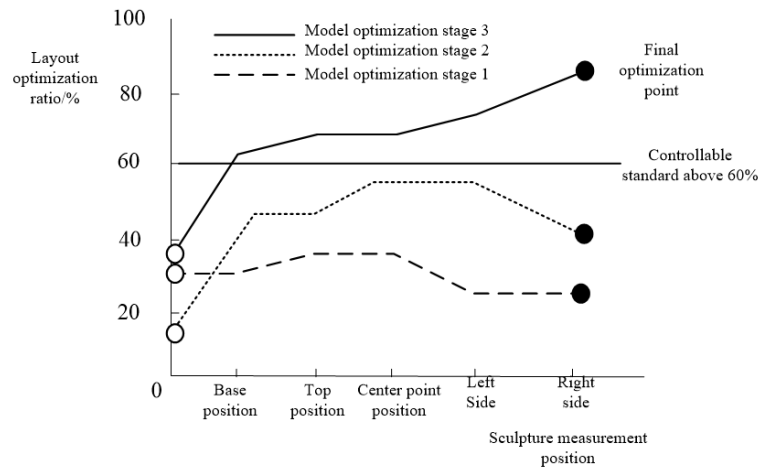


Fig. 3.2: Comparison and Analysis of Test Results

significant potential for expanding the space dedicated to urban sculpture. By employing a multi-level approach to planning control and considering local conditions, the model guides the characteristics, culture, and even the setting environment of the city to design specific planning schemes. Furthermore, it emphasizes the meticulous multi-dimensional processing of details and incorporates hierarchical control, placing a strong emphasis on the key features of sculpture. Through this comprehensive design approach, the research achieves multi-dimensional planning and control of space, effectively integrating multiple objectives into the landscape ceramic sculpture design process. Looking ahead, this work opens avenues for further innovation and development of the overall layout planning model for sculptures, taking it to new heights of creativity and effectiveness. Future research can explore additional optimization techniques, incorporate advanced technologies, and consider evolving urban design principles to continue refining the field of landscape ceramic sculpture planning.

#### REFERENCES

- [1] D. ATHUKORALA, R. C. ESTOQUE, Y. MURAYAMA, AND B. MATSUSHITA, *Impacts of urbanization on the mthurajawela marsh and negombo lagoon, sri lanka: Implications for landscape planning towards a sustainable urban wetland ecosystem*, *Remote Sensing*, 13 (2021), p. 316.
- [2] H. CHOE, A. T. KEELEY, D. R. CAMERON, M. GOGOL-PROKURAT, L. HANNAH, P. R. ROEHRDANZ, C. A. SCHLOSS, AND J. H. THORNE, *The influence of model frameworks in spatial planning of regional climate-adaptive connectivity for conservation planning*, *Landscape and Urban Planning*, 214 (2021), p. 104169.
- [3] P. CSORBA, *Present tendencies in landscape planning and recognition of the opinion of local citizens on the example of the tiszazug*, *Acta geographica Debrecina Landscape & Environment series*, 15 (2021), pp. 1–9.
- [4] G. ENEMAN, A. VELOSO, P. FAVIA, A. HIKAVYY, C. PORRET, H. ARIMURA, A. DE KEERSGIETER, G. POURTOIS, R. LOO, A. SPESSOT, ET AL., *Stress in silicon-germanium nanowires: Layout dependence and imperfect source/drain epitaxial stressors*, *IEEE Transactions on Electron Devices*, 68 (2021), pp. 5380–5385.
- [5] L. ERNOUL, A. WARDELL-JOHNSON, R. MATHEVET, A. SANDOZ, O. BOUTRON, L. WILLM, S. ARNASSANT, AND A. BÉCHET, *Context in landscape planning: Improving conservation outcomes by identifying social values for a flagship species*, *Sustainability*, 13 (2021), p. 6827.
- [6] S. GUO AND B. WANG, *Application of computer aided modeling design in the expression techniques of sculpture art space*, *Computer-Aided Design and Applications*, 19 (2021), pp. 1–12.
- [7] J. LI AND T. HOU, *Application of virtual reality technology in analysis of the three-dimensional evaluation system of rural landscape planning*, *Mathematical Problems in Engineering*, 2021 (2021), pp. 1–16.
- [8] Q. LI AND P. WANG, *River wetland landscape planning and design from low-carbon perspective.*, *Arabian Journal of Geosciences*, 14 (2021).
- [9] S. LI AND Z. FAN, *Evaluation of urban green space landscape planning scheme based on pso-bp neural network model*, *Alexandria Engineering Journal*, 61 (2022), pp. 7141–7153.
- [10] L. LIANG, *Optimal planning simulation of color balance layout of urban ecological landscape*, *Computer Simulation*, 38 (2021), p. 4.

- [11] J. LIU, Q. CHEN, AND X. TIAN, *Illustration design model with clustering optimization genetic algorithm*, Complexity, 2021 (2021), pp. 1–10.
- [12] E. MESELHE, E. WHITE, Y. WANG, AND D. REED, *Uncertainty analysis for landscape models used for coastal planning*, Estuarine, Coastal and Shelf Science, 256 (2021), p. 107371.
- [13] S. MEYER, *Implied dynamics: The ceramic sculpture of kenneth baskin*, Ceramics Monthly, (2021), p. 69.
- [14] D. MIKIĆ, H. OTMAČIĆ ČURKOVIĆ, T. KOSEC, AND N. PEKO, *An electrochemical and spectroscopic study of surfaces on bronze sculptures exposed to urban environment*, Materials, 14 (2021), p. 2063.
- [15] S. PENA, M. ABREU, AND M. MAGALHÃES, *Rethinking coastal cliff protection zones for landscape planning. what limits are enough?*, Applied Geography, 127 (2021), p. 102387.
- [16] L. PENG, *Intelligent landscape design and land planning based on neural network and wireless sensor network*, Journal of Intelligent & Fuzzy Systems, 40 (2021), pp. 2055–2067.
- [17] L. PENG, L. YU, H. SHEN, AND J. PI, *3d garden landscape planning visualization system based on fpga processor and virtual reality*, Microprocessors and Microsystems, 81 (2021), p. 103698.
- [18] B. W. PENNAROLA, A. FRY, L. PRICHETT, A. E. BERI, N. N. SHAH, AND L. WIENER, *Mapping the landscape of advance care planning in adolescents and young adults receiving allogeneic hematopoietic stem cell transplantation: A 5-year retrospective review*, Transplantation and Cellular Therapy, 28 (2022), pp. 164–e1.
- [19] R. RYAN, *Selected papers from the 6th fábos conference on landscape and greenway planning: Adapting to expanding and contracting cities ed. by richard smardon*, Landscape Journal: design, planning, and management of the land, 40 (2021), pp. 115–118.
- [20] Y. SAHRAOUI, C. D. G. LESKI, M.-L. BENOT, F. REVERS, D. SALLES, I. VAN HALDER, M. BARNEIX, AND L. CARASSOU, *Integrating ecological networks modelling in a participatory approach for assessing impacts of planning scenarios on landscape connectivity*, Landscape and Urban Planning, 209 (2021), p. 104039.
- [21] P. SHAN AND W. SUN, *Research on landscape design system based on 3d virtual reality and image processing technology*, Ecological Informatics, 63 (2021), p. 101287.
- [22] J. WAN, F. ZHANG, AND J. PAN, *Promoting organized basic research: Strategic layout and strategic capacity in science and technology*, Bulletin of Chinese Academy of Sciences (Chinese Version), 36 (2021), pp. 1404–1412.
- [23] F. WITING, M. A. E. FORIO, F. J. BURDON, B. MCKIE, P. GOETHALS, M. STRAUCH, AND M. VOLK, *Riparian reforestation on the landscape scale: Navigating trade-offs among agricultural production, ecosystem functioning and biodiversity*, Journal of Applied Ecology, 59 (2022), pp. 1456–1471.
- [24] Y. XIANXIAN, N. CHANGHUIZI, B. YIFEI, AND Y. SHENGYUAN, *Application of ecological and environmental protection concept in urban landscape planning and design*, Journal of Environmental Protection and Ecology, 22 (2021), pp. 2693–2700.

*Edited by:* Sathishkumar V E

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

*Received:* May 12, 2023

*Accepted:* Jul 9, 2023



## EVALUATION OF CURRICULUM IMPLEMENTATION EFFECTIVENESS OF HIGHER VOCATIONAL EDUCATION BASED ON COLLABORATIVE FILTERING ALGORITHM

**Abstract.** The conventional evaluation method for the effectiveness of curriculum implementation mainly focuses on the complete orientation analysis of students' curriculum content, which does not reflect the value of the educational curriculum and affects the effectiveness of evaluation. Therefore, an evaluation method of curriculum implementation effectiveness of higher vocational education based on a collaborative filtering algorithm is designed. Identify the practical focus of evaluating the implementation of the higher vocational education curriculum and discover the educational curriculum's significance. Qualitative evaluation of curriculum implementation degree based on collaborative filtering algorithm, find out the hidden characteristics of curriculum implementation evaluation to effectively evaluate higher vocational education curriculum implementation. Using case analysis, it is verified that the method is more effective and can be applied in real life.

**Key words:** Collaborative filtering algorithm; Higher vocational education; Curriculum implementation; Effectiveness; Evaluation method;

**1. Introduction.** The educational theory is constantly changing. The scientific demonstration effect of its logic theory is concentrated in the aspect of curriculum evaluation, which provides a scientific orientation for teaching curriculum theory. The promotion of professional courses is related to many variables such as students' learning ability, teachers' professional ability, students' professional skills, and learning process [11, 22, 5].

Education and training are prerequisites for students to recognize and enter society. Adhering to the concept of "natural education", it is the curriculum learning process for students. It sets students' self-learning and self-evaluation function for ensuring they have specific knowledge and skills. Scientific curriculum implementation evaluation contains basic assumptions. Among the total multiple evaluation groups, only minority groups are outstanding, and most of them are at normal level [15, 8].

The motivation behind the work lies in recognizing the limitations of conventional evaluation methods for assessing the effectiveness of curriculum implementation in higher vocational education. The existing approaches primarily focus on analyzing the completion and orientation of students' curriculum content, often neglecting the value and significance of educational curriculum. This limitation hinders the accurate evaluation of curriculum implementation and compromises the quality of teaching in higher vocational education.

To address this gap, the authors of this work were motivated to propose a novel evaluation method based on the collaborative filtering algorithm. This algorithm, commonly used in recommendation systems, offers a fresh perspective on evaluating curriculum implementation effectiveness. By leveraging collaborative filtering, the method aims to identify the practical focus of evaluating curriculum implementation, uncover the hidden characteristics of the evaluation process, and provide a more comprehensive assessment of higher vocational education curriculum effectiveness.

The motivation also stems from the desire to improve teaching quality in higher vocational education. The authors aim to enhance accountability, ensure quality assurance, and promote continuous improvement in teaching practices by developing a more effective evaluation method. The evaluation method provides educators with valuable feedback and insights to reflect on their teaching strategies, make evidence-based decisions, and tailor instruction to meet the diverse needs of students.

The evaluation of curriculum implementation is to select a few outstanding people, let them experience the happiness of success, and promote the exceptional people to be more outstanding using incentives. Throughout the evaluation procedure, it is essential to guarantee fairness and objectivity in the assessment, it is necessary to make the evaluation value neutral, and the evaluated cannot participate in the evaluation process to ensure the accuracy of the evaluation [16, 3, 12]. This article will focus on evaluating curriculum implementation and finding the best curriculum implementation plan to improve teaching quality. To conduct a more precise assess-

ment of the efficacy of curriculum implementation in higher vocational colleges, this paper uses a collaborative filtering algorithm to design an evaluation method for the effectiveness of curriculum implementation in higher vocational education.

The work described proposes a novel evaluation method for assessing the effectiveness of curriculum implementation in higher vocational education. The conventional evaluation methods typically focus on analyzing the orientation and completion of students' curriculum content, which fails to capture the true value of educational curriculum and consequently impacts the accuracy of the evaluation.

The authors designed an evaluation method based on the collaborative filtering algorithm to address this limitation. This approach aims to identify the practical focus of evaluating the implementation of higher vocational education curricula and uncover the significance of educational curricula. By employing qualitative evaluation techniques, the method assesses the degree of curriculum implementation based on the collaborative filtering algorithm, allowing for discovering hidden characteristics in the evaluation process.

The main contributions of this work can be summarized as follows:

1. The work introduces a new evaluation method that considers the collaborative filtering algorithm, providing a fresh perspective on assessing the effectiveness of curriculum implementation in higher vocational education.
2. The proposed method emphasizes identifying the practical aspects of evaluating the implementation of higher vocational education curricula. This allows for a more targeted evaluation process that aligns with the real-life requirements of vocational education.
3. By utilizing the collaborative filtering algorithm, the method uncovers the true significance of the educational curriculum. This helps in understanding the value and impact of curriculum implementation on students' learning outcomes.
4. The work incorporates qualitative evaluation techniques to assess the degree of curriculum implementation. This approach provides a deeper understanding of the evaluation process and facilitates the identification of hidden characteristics that influence effectiveness.
5. Through case analysis, the authors validate the effectiveness of the proposed method and demonstrate its applicability in real-life scenarios. This contributes to bridging the gap between theory and practice in evaluating the implementation of higher vocational education curriculum enumerate

The design of the evaluation method for curriculum implementation effectiveness in higher vocational education, based on the collaborative filtering algorithm, is a key contribution of this work. This innovative method aims to overcome the limitations of conventional evaluation approaches that primarily focus on the analysis of students' curriculum content orientation.

The key steps involved in designing the evaluation method are as follows:

1. The method leverages the collaborative filtering algorithm, which is widely used in recommendation systems, to evaluate curriculum implementation effectiveness. Collaborative filtering analyzes patterns of student behavior and preferences to make predictions and recommendations.
2. The evaluation method identifies the practical focus for evaluating curriculum implementation in higher vocational education. This involves understanding the curriculum's specific objectives, requirements, and outcomes and aligning the evaluation criteria accordingly.
3. The method incorporates qualitative evaluation techniques to assess the degree of curriculum implementation. This may involve gathering feedback from teachers, students, and other stakeholders through surveys, interviews, or observations. The qualitative evaluation helps uncover hidden characteristics and subjective aspects of curriculum implementation.
4. By employing the collaborative filtering algorithm, the method explores the significance of the educational curriculum. It seeks to understand how the curriculum contributes to students' learning outcomes, skill development, and overall educational experience.
5. The evaluation method aims to identify hidden characteristics of curriculum implementation evaluation. This involves analyzing various factors that influence the effectiveness of curriculum implementation, such as teaching methods, instructional materials, assessment strategies, and student support services. The collaborative filtering algorithm helps reveal patterns and relationships among these characteristics.
6. The evaluation method measures the effectiveness of curriculum implementation based on the insights

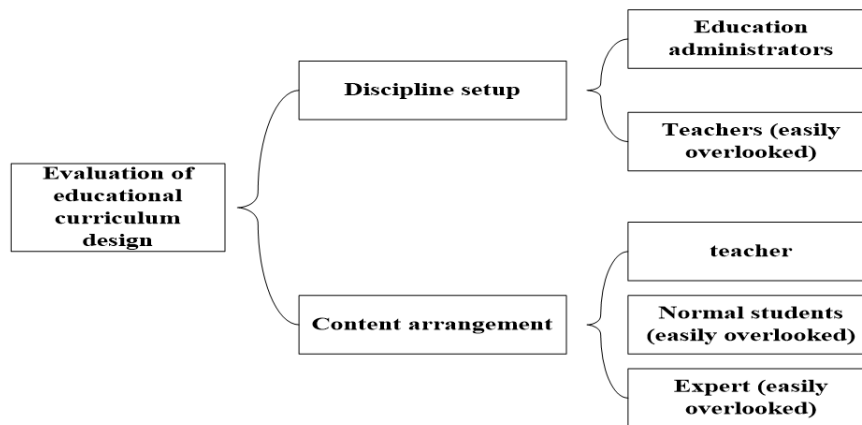


Fig. 1.1: Overview of Education Curriculum Design Evaluation

obtained through collaborative filtering and qualitative evaluation. It provides a comprehensive assessment of how well the curriculum is implemented, highlighting strengths, weaknesses, and areas for improvement.

7. The proposed method is validated through case analysis. Real-life examples of higher vocational education curriculum implementation are examined to demonstrate the effectiveness and practical application of the evaluation method. This step ensures that the method is theoretically sound and applicable in real-world contexts.

**1.2. Determine the functional orientation of curriculum implementation evaluation of higher vocational education.** The evaluation of educational curriculum implementation evaluates the value of a teacher education curriculum. The assessment of the design, implementation, and impact of teacher education curriculum is the key aspect of curriculum evaluation in higher vocational education [24, 14, 17]. To assess the effectiveness of curriculum implementation, it is essential to establish a clear agenda for evaluation and analyze the planning process of the educational curriculum [25, 9, 23]. The evaluation overview is shown in Figure 1.1.

As shown in Figure 1.1, education curriculum design is the standard for evaluating the implementation of teaching curriculum, and the evaluation standard is the criterion for value judgment in the implementation evaluation process. Therefore, reviewing educational curriculum design is the primary function of curriculum implementation evaluation. In the evaluation process, the principles of appropriateness, effectiveness, feasibility, and accuracy are always followed [4, 18, 7]. The evaluation of course implementation reflects the fundamental characteristics of professional practice courses in higher vocational colleges, and the ideal evaluation grade is obtained with the expected evaluation standard to improve the overall quality of teaching.

**1.3. Qualitative evaluation of course implementation based on collaborative filtering algorithm.** In curriculum implementation the orientation of curriculum implementation takes many forms. No matter which form it takes, it will affect the effectiveness of evaluation [1]. This paper analyzes the evaluation orientation of curriculum implementation in higher education settings from three perspectives of technology, politics, and culture, as shown in Table 1.1.

As shown in Table 1.1, this paper regards curriculum implementation evaluation as a kind of technology from the perspective of technology. The curriculum implementation evaluation is the process of implementing the scheduled plan, and the measurement standard is the degree of achievement of teaching objectives. From the political point of view, the curriculum implementation evaluation is regarded as the interest competition between groups, and different competitive relationships maintain different attitudes towards the curriculum implementation evaluation [2, 6, 20]. From the culture perspective, curriculum implementation evaluation

Table 1.1: Curriculum implementation orientation of higher vocational education

Implementation orientation	Fidelity orientation Technological outlook	Mutual adaptation orientation Political outlook	Curriculum Creation Orientation Cultural outlook
Basic assumptions	Implementation is a technical task, and the key lies in finding the best means to achieve curriculum evaluation	Implementation is a two-way social interaction process, and consensus is reached through implementation orientation evaluation	Implementation is a non-linear and complex evaluation evolution process
Evaluation focus	The rationality of educational curriculum implementation; Degree and efficiency of curriculum implementation	Interactions between school scenarios and evaluation plans to adapt the implementation of the evaluation process	The content and impact of school context and creation
Implementation evaluation	Application of specialized knowledge; RD&D mode	Flexible evaluation; RAND mode	Grassroots evaluation, TORL mode
Evaluation method	Quantitative evaluation	Quantitative+qualitative evaluation	Qualitative evaluation

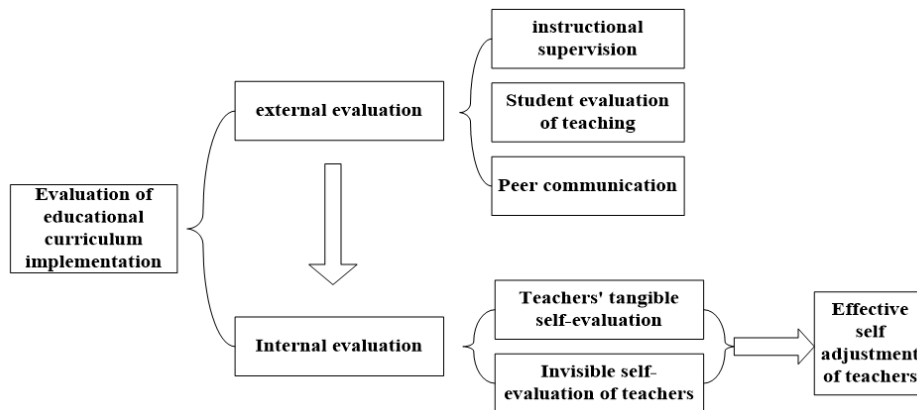


Fig. 1.2: Overview of Education Curriculum Implementation Evaluation

is regarded as cultural regeneration, and curriculum implementation evaluation is the process of promoting teachers to think over teaching arrangements. Analyzing the impact of curriculum implementation assessment from diverse viewpoints holds significance in enhancing the efficacy of curriculum implementation evaluation [19]. An overview of the implementation evaluation of educational courses is shown in Figure 1.2.

As shown in Figure 1.3, the assessment of the implementation of the curriculum in higher vocational education is predominantly grounded on the merits of the content and the worth of self-reflection, which empowers instructors to leverage their pedagogical aptitude and students to reinforce their cognitive competence [13]. This paper employs the collaborative filtering algorithm to enhance the effectiveness of evaluating curriculum implementation in higher vocational education, which analyzes similarities among curricula. The formula is as

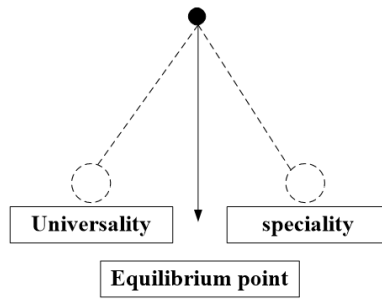


Fig. 1.3: Curriculum Orientation of Higher Vocational Education

follows:

$$sim(a, b) = \cos(a, b) = \frac{a * b}{\|a\| * \|b\|} \tag{1.1}$$

In formula 1.1,  $sim(a, b)$  is the similarity vector of the courses that students like;  $\cos(a, b)$  is cosine similarity;  $a, b$  two students;  $\|a\| * \|b\|$  in the position of denominator, the smaller the product, the student  $a$  with students  $b$  the higher the similarity of the courses they like. Putting students  $a$  with students  $b$  learning in the same class can improve the degree of curriculum implementation. As far as the collaborative filtering algorithm is concerned, the relationship between the two variables is used to weigh the similarity of students' preferences for the curriculum to provide students with a good curriculum implementation environment [10]. According to students  $a$  with students  $b$  the linear correlation of  $sim(a, b)$  the value range of is set between - 1 and 1, and the linear correlation formula is:

$$sim(a, b)' = \frac{\sum_i (R_{a,i} - \bar{R})(R_{b,i} - \bar{R})}{\sqrt{\sum_{i \in I} (R_{a,i} - \bar{R})^2} \sqrt{\sum_{i \in I} (R_{b,i} - \bar{R})^2}} \tag{1.2}$$

In formula 1.2,  $sim(a, b)'$  for the similarity of students' interest in the process of curriculum implementation;  $R_{a,i}$  for students  $a$  for the  $i$  evaluation of the performance of the courses;  $R_{b,i}$  for students  $b$  for the  $i$  evaluation of the implementation of the courses;  $\bar{R}$  is an index for evaluating the balance. After analyzing the degree of students' interest in curriculum content in the course of curriculum implementation, this paper establishes adjacent sets. Hypothetical students  $a$  the adjacent set of is  $N_a$ , students  $a$  be interested in a course  $i$  the estimated predictive value of is  $P_{a,i}$   $a$  the formula for evaluating the implementation of courses in the neighborhood set of is as follows:

$$P_{a,i} = \bar{R} + \frac{\sum_{n \in N_a} sim(a, n)' * (R_{n,i} - \bar{R})}{\sum_{n \in N_a} (|sim(a, n)'|)} \tag{1.3}$$

In equation 1.3,  $P_{a,i}$  for students  $a$  be interested in a course  $i$  the predicted estimated value of;  $sim(a, n)'$  for new students  $n$  with students  $a$  satisfactory similarity of curriculum implementation;  $R_{n,i}$  for students  $n$  for the  $i$  evaluation of the implementation of the courses. When applying the collaborative filtering algorithm to evaluate the performance of a higher vocational education curriculum, analyzing the nearest neighbor of the target course is critical. In the neighbor evaluation step, find the neighbor set of the target curriculum and measure the implementation similarity of each curriculum [21]. The hidden characteristics of curriculum implementation are analyzed, as shown in Figure 1.3.

As shown in Figure 1.3, this paper finds out the hidden features of curriculum implementation by analyzing the degree of interest of students in the curriculum and reduces the dimensions of more complex student evaluation content to obtain simple and easy-to-operate hidden features. The matrix decomposition expression

is:

$$\bar{R}_{m \times n} = P^t_{m \times k} \times Q_{k \times n} \quad (1.4)$$

In equation 1.4,  $\bar{R}_{m \times n}$  original scoring matrix for  $m$  students to implement  $n$  courses;  $P^t_{m \times k}$  find out  $k$  hidden features for  $m$  students and then evaluate the  $t$ -th education curriculum;  $Q_{k \times n}$  integrate the score items of matrix  $Q$  with  $k$  hidden features for  $n$  courses. In order to ensure the effectiveness of education curriculum implementation evaluation, this paper uses the loss function to reduce evaluation error and adds regularization terms to prevent overfitting. The loss function is expressed as follows:

$$V = \sum_{(a,i) \in R} (R_{ai} - P^t_{m \times k} - Q_i)^2 + \gamma \sum_a \|P^t_{m \times k}\|^2 + \gamma \sum_i \|Q_i\|^2 \quad (1.5)$$

In formula 1.5,  $V$  is the loss function expression;  $R_{ai}$  is the evaluation error;  $Q_i$  is a regular scoring item;  $\gamma$  is a cross parameter. This paper determines the feasible implementation effectiveness  $d$ , so that it will decline along the fastest defense line gradient, then:

$$P_a = P_a + d(Q_i - \gamma P_a) \quad (1.6)$$

In formula 1.6,  $P_a$  for students  $a$  the evaluation and optimization indicators of curriculum implementation. Collaborative filtering algorithms can be used to improve the accuracy of student evaluations by identifying hidden features of curriculum implementation. By analyzing previous evaluations and identifying patterns, collaborative filtering algorithms can find relationships between different aspects of curriculum implementation and student ratings. This can assist in pinpointing regions that require enhancement and give an understanding into which elements are crucial in evaluating the execution of a curriculum.

Collaborative filtering algorithm for higher vocational education:

1. Input:
  - (a) Curriculum data (courses, teaching materials, assessments, etc.)
  - (b) User data (students' historical behavior, preferences, feedback)
  - (c) Evaluation criteria and metrics
2. Data Preprocessing:
  - (a) Normalize user preferences and curriculum data to ensure consistency and comparability.
  - (b) Handle missing data, if any, through imputation techniques.
3. User-Based Collaborative Filtering:
  - (a) Compute the similarity between target users and other users based on their curriculum-related behavior and preferences (e.g., courses taken, grades obtained, feedback provided).
  - (b) Select a set of similar users based on similarity threshold or top-K nearest neighbors.
4. Item-Based Collaborative Filtering:
  - (a) Compute the similarity between target curriculum components (e.g., courses, teaching materials) based on user preferences and feedback.
  - (b) Select a set of similar curriculum components based on similarity threshold or top-K nearest neighbors.
5. Predicting Effectiveness:
  - (a) For user-based collaborative filtering:
    - i. Aggregate the curriculum-related feedback and ratings of similar users.
    - ii. Predict the effectiveness of curriculum components for the target user based on the aggregated feedback.
  - (b) For item-based collaborative filtering:
    - i. Aggregate the user preferences and feedback for similar curriculum components.
    - ii. Predict the effectiveness of the target curriculum component based on the aggregated feedback.
6. Evaluation:
  - (a) Compare the predicted effectiveness with the predefined evaluation criteria and metrics.



Table 2.1: Talent Training Specifications of M Higher Vocational Colleges

Type	Talent training specifications
Quality requirements	Love national culture; Abide by laws and regulations, school rules and regulations; Have good professional ethics and innovative practical spirit; Having aesthetic taste, sound psychology, and healthy physique; Have solid professional knowledge and ability; Have a good sense of service and team spirit; Strong communication and organizational skills
Knowledge requirements	Master comprehensive knowledge of listening, speaking, reading, and writing; Proficient in computer knowledge; Master professional theories on installation, commissioning, maintenance, repair, and management of electromechanical equipment; Familiar with mechanical manufacturing theory; Master electrical technology operation, etc
Capability requirements	Basic Chinese language skills and written expression skills; Ability to analyze and troubleshoot mechanical structures; Ability to read and draw mechanical drawings; Ability to detect and maintain electrical equipment; Equipped with basic bench work operations; Ability to control, install, debug, and operate common electrical equipment

- (b) Assess the degree of curriculum implementation effectiveness based on the evaluation results.
- 7. Output:
  - (a) Evaluation results indicating the effectiveness of curriculum implementation for different curriculum components.
  - (b) Personalized recommendations for students based on their preferences and predicted effectiveness.

**2. Example analysis.**

**2.1. Overview of higher vocational colleges.** In order to verify whether the curriculum implementation evaluation method designed in this paper can be applied to real life, this article utilizes M vocational college as a case study to examine the aforementioned approaches. With the teaching characteristics of serving the rail transit industry, the teaching theme of serving the equipment manufacturing industry, and the form of "double closed-loop control", it established a teaching quality assurance system with local characteristics, providing employment security for students in the college. With the popularization of online education mode, M Higher Vocational College has unified the teaching content, occupation, competition, and assessment standards, and combined the requirements for students with enterprises. It has formed more than 50 social organizations such as the "assistant group", "reporter group", and "military band", providing students with rich spare time life. The talent training specifications of higher vocational colleges are shown in Table 2.1.

As shown in Table 2.1, M higher vocational college is an automobile repair college, mainly setting up electromechanical equipment maintenance and management specialty, electromechanical integration technology, etc., to create strong practical ability and innovation promotion for students. Provide the market with advanced technical applied talents from the installation, commissioning, maintenance, repair, sales, management, and other aspects of electromechanical equipment. Colleges and universities mainly cultivate students' comprehensive ability from the aspects of quality requirements, knowledge requirements, ability requirements, etc. From the student level, establishing a correct concept of learning can become, promote students to develop learning habits. In order to improve the learning effect of students, and based on the identified relationships, the algorithm can provide insights into the effectiveness of the implementation of educational courses and suggest improvements if needed. An example of collaborative filtering evaluation is shown in Figure 2.1.

As shown in Figure 2.1, student A likes courses A and C; Student B likes course B; Student C likes courses A, C, and D. In this paper, student C is regarded as the neighbor of student A, and can be recommended to student A what student A has not learned, so as to ensure that students can recommend courses to each other, which is conducive to the implementation of courses. Evaluate the implementation of educational courses

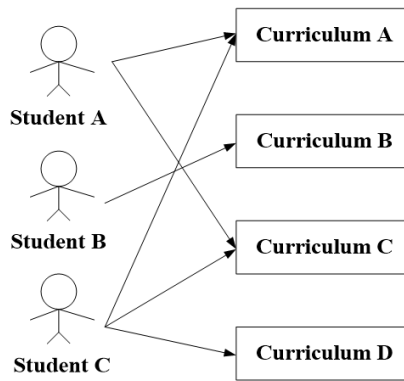


Fig. 2.1: Example of collaborative filtering evaluation

Table 2.2: Application Results

Variable	sexual distinction	mean value	standard deviation	T value	P value
Course Functions	man	4.160	0.615	-3.358	0.001
	woman	4.279	0.514		
Course structure	man	3.814	0.830	1.044	0.096
	woman	3.764	0.780		
Course content	man	3.858	0.782	2.351	0.068
	woman	3.749	0.771		
Teaching methods	man	4.135	0.621	-3.650	0.000
	woman	4.263	0.574		
Course evaluation	man	4.030	0.735	-2.217	0.027
	woman	4.125	0.656		
course management	man	3.801	0.985	0.125	0.090
	woman	3.809	0.985		

according to student’s preference for the courses to ensure the effectiveness of the implementation evaluation.

**2.2. Application results.** Under the above conditions, this paper randomly selects variables such as curriculum function, structure, content, teaching methods, evaluation, management, etc.. It analyzes the differences in curriculum implementation degrees between men and women under different gender conditions. The application results evaluate the effectiveness of curriculum implementation based on the change of T value and P value. The application results are shown in Table 2.2.

As shown in Table 2.2, teachers of different sexes have different degrees of implementation in terms of curriculum functions, teaching methods, curriculum evaluation, etc., and the T value is negative. According to the standard deviation, female teachers are significantly higher than male teachers in terms of curriculum functions, teaching methods, curriculum evaluation, etc. The P value shows the evaluation result of the degree of curriculum implementation. The smaller the P value, the higher the evaluation effectiveness of educational curriculum implementation; The higher the P value, the lower the effectiveness of the evaluation of educational curriculum implementation. It can be seen from the table that after using the effectiveness evaluation method designed in this paper, the P value < 0.1, the evaluation effectiveness of the implementation of educational courses is relatively high, and the education situation of M vocational colleges can be clearly analyzed.

**3. Conclusion.** In recent years, higher vocational colleges have emerged as crucial institutions for providing practical education and training to students. It is imperative to continually enhance curriculum implementation and evaluation methods to ensure students receive the best possible education. The evaluation of curriculum implementation in higher vocational institutions often involves a combination of qualitative and quantitative methods. These evaluation standards consider various factors that influence the quality of education, aiming to enhance the effectiveness of curriculum implementation and foster value-added educational outcomes. Higher vocational education plays a vital role in cultivating employment-oriented talents, with professional practice courses serving as essential conduits for practical skill development. The evaluation of curriculum implementation directly influences the degree to which talent training objectives are achieved, significantly impacting students' future employability. To further enhance students' learning outcomes, this study utilizes the collaborative filtering algorithm to design an evaluation method for assessing the effectiveness of curriculum implementation in higher vocational education. It is important to emphasize that curriculum evaluation should be an ongoing process, continuously adapting to meet the evolving needs of industries and society as a whole. This includes evaluating the degree of curriculum implementation and addressing other relevant aspects. By engaging in continuous evaluation, higher vocational institutions can ensure that their curriculum remains relevant and effective, providing students with a solid foundation for their future employment.

## REFERENCES

- [1] A. ALMU AND Z. BELLO, *An experimental study on the accuracy and efficiency of some similarity measures for collaborative filtering recommender systems*, (2021).
- [2] L. BADIS, M. AMAD, D. AÏSSANI, AND S. ABBAR, *P2pcf: a collaborative filtering based recommender system for peer to peer social networks*, Journal of High Speed Networks, 27 (2021), pp. 13–31.
- [3] J. BIAN, *Research on aesthetic thinking and computer courseware design of higher vocational aesthetic education based on information perspective*, in Journal of Physics: Conference Series, vol. 1744, IOP Publishing, 2021, p. 032109.
- [4] O. DENISOVA, G. KUNSBAEVA, AND A. CHIGLINTSIVA, *Big data: some ways to solve the problems of higher education*, in Journal of Physics: Conference Series, vol. 2001, IOP Publishing, 2021, p. 012021.
- [5] L. GRAF AND A. P. LOHSE, *Conditions for cross-border policy transfer and cooperation: Analysing differences between higher education and vocational training*, Research in Comparative and International Education, 16 (2021), pp. 361–383.
- [6] P.-Y. HSU, J.-Y. CHUNG, AND Y.-C. LIU, *Using the beta distribution technique to detect attacked items from collaborative filtering*, Intelligent Data Analysis, 25 (2021), pp. 121–137.
- [7] C. IWENDI, E. IBEKE, H. EGGONI, S. VELAGALA, AND G. SRIVASTAVA, *Pointer-based item-to-item collaborative filtering recommendation system using a machine learning model*, International Journal of Information Technology & Decision Making, 21 (2022), pp. 463–484.
- [8] T. D. KARMINSKAYA AND V. F. ISLAMUTDINOV, *Influence of higher and vocational education on the economic development of the khmao-yugra region*, Ekonomika Regiona= Economy of Regions, (2021), p. 445.
- [9] S. KARSTINA, O. ZECHIEL, AND C. MACHADO, *Role of the kazakhstan-german cooperation in improving scientific tools for evaluation of vocational education programs*, Vysshee Obrazovanie v Rossii= Higher Education in Russia, 30 (2021), pp. 132–143.
- [10] T.-Y. KIM, H. KO, S.-H. KIM, AND H.-D. KIM, *Modeling of recommendation system based on emotional information and collaborative filtering*, Sensors, 21 (2021), p. 1997.
- [11] S.-Y. LEE, J. C.-K. LEE, AND B. Y.-H. LAM, *Does renaming improve public attitudes toward vocational education and training in higher education? evidence from a survey experiment*, Education+ Training, 64 (2022), pp. 347–359.
- [12] X. LI, *Current situation and measures to improve the ideological and political education among students in higher vocational institutions*, Journal of Contemporary Educational Research, 5 (2021), pp. 96–99.
- [13] D. LIAN, J. CHEN, K. ZHENG, E. CHEN, AND X. ZHOU, *Ranking-based implicit regularization for one-class collaborative filtering*, IEEE Transactions on Knowledge and Data Engineering, 34 (2021), pp. 5951–5963.
- [14] B. LIN, *Importance of ideological and political education in teaching fine arts education in higher vocational colleges*, Journal of Contemporary Educational Research, 5 (2021), pp. 86–89.
- [15] W. LIN, *Innovative thinking in higher vocational colleges' physical education*, Journal of Contemporary Educational Research, 5 (2021), pp. 14–18.
- [16] H. A. MAULANA, *Psychological impact of online learning during the covid-19 pandemic: A case study on vocational higher education*, Indonesian Journal of Learning Education and Counseling, 3 (2021), pp. 130–139.
- [17] B. MÁTÉ-SZABÓ AND D. A. TÓTH, *Appearance and development of short-term higher education vocational training in hungary*, Acta Educationis Generalis, 11 (2021), pp. 63–76.
- [18] S. POUDEL AND M. BIKDASH, *Optimal dependence of performance and efficiency of collaborative filtering on random stratified subsampling*, Big Data Mining and Analytics, 5 (2022), pp. 192–205.
- [19] R. WANG AND T. LI, *Collaborative filtering recommendation algorithm based on user behavior and drug semantics*, in Journal of Physics: Conference Series, vol. 1802, IOP Publishing, 2021, p. 032005.
- [20] X. WANG, Z. DAI, H. LI, AND J. YANG, *Research on hybrid collaborative filtering recommendation algorithm based on the*

- time effect and sentiment analysis*, Complexity, 2021 (2021), pp. 1–11.
- [21] Z. R. WANG, S. W. SONG, AND H. C. YAN, *Performance comparison and analysis of collaborative filtering recommendation algorithms*, Computer Simulation, 39 (2022), pp. 435–440.
- [22] J. WU, *Research on computer basic teaching in higher vocational education under the background of flipped classroom*, in Journal of Physics: Conference Series, vol. 1738, IOP Publishing, 2021, p. 012124.
- [23] S. YAMADA AND C. S. OTCHIA, *Perception gaps on employable skills between technical and vocational education and training (tvvet) teachers and students: the case of the garment sector in ethiopia*, Higher Education, Skills and Work-Based Learning, 11 (2021), pp. 199–213.
- [24] Y. YU, *Based on the apartment community management mode to explore the education strategy of higher vocational colleges—take polus international college as an example*, Open Journal of Social Sciences, 9 (2021), pp. 322–333.
- [25] Y. ZHOU, *The application of curriculum ideology and politics in the training of judicial vocational education talents*, Journal of Higher Education Research, 3 (2022), pp. 155–159.

*Edited by:* Sathishkumar V E

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

*Received:* May 12, 2023

*Accepted:* Jul 10, 2023



## A METHOD FOR ONLINE MONITORING DATA RELEASE OF COMPOSITE SUBMARINE CABLE BASED ON HORIZONTAL FEDERATED LEARNING

**Abstract.** Conventional online composite submarine cable monitoring data release mostly adopts the method and principle of blockchain dynamic zoning consensus. In the data release process, there are omissions, and it takes a long time to complete the task, which reduces the timeliness of online composite submarine cable monitoring data release. Based on this, a new data publishing method is proposed by introducing horizontal federation learning. First, the online monitoring data of composite submarine cables are collected and preprocessed to eliminate the high-frequency capacitive effect of submarine cables. Secondly, manage composite submarine cable data nodes, transform the status relationship of data nodes, and ensure the quality of subsequent data release. A horizontal federation learning model is established to design the online monitoring data release process. The experimental results show that the new data release method is highly feasible. With the increasing online monitoring data of composite submarine cables, the time required for data release is short, and the timeliness is high.

**Key words:** Horizontal federal learning; Reunite with; Monitor; Submarine cable; On-line; Data; Release;

**1. Introduction.** With the internationalization of power grid, the development and utilization of offshore oil platforms and offshore wind power, and the rapid development of island economy, the number of submarine cables has increased rapidly [3]. To solve the problems of dense power grid structure and unbalanced power generation structure, submarine cable projects have been carried out at home and abroad. Since 1990, the total length of submarine cable projects has been 13618km, with a full design capacity of 35618MW [6]. As the number of cables increases, many problems begin to surface. Compared with land cables, submarine cables are difficult to find faults, and the repair time is long, and the repair cost is expensive [8]. The submarine cable laying environment is complex and vulnerable to external forces such as anchoring and fishing [4]. When laying cables, it is easy to leave some defects, such as a damaged outer protective layer, armor layer and insulation layer [17], due to the influence of sea waves, seabed geology, personnel, and equipment.

The length of submarine cable varies from several kilometers to tens of kilometers, and there is no obvious cable route on the sea. It is impossible to accurately and timely detect the ships entering the cable reserve by relying on manual patrol and lookout [18]. Submarine cables are relatively heavy, requiring large engineering ships and professionals to lay and maintain them. The complexity of the submarine cable structure and the marine environment, ocean current, earthquake, and ship anchoring can not accurately and timely judge the damage or degree of damage caused to the cable after operation, and in turn, can not determine whether maintenance is required [20]. The routine test of power failure and maintenance of submarine cables is a loss of power and an additional investment. The online monitoring research of composite submarine cables is proposed to ensure the safe operation of submarine cables. According to the online monitoring data, the operating conditions of submarine cables are obtained, and the potential safety risks of submarine cables [22] are found in time. After the online monitoring of composite submarine cables is completed, the online monitoring data of submarine cables need to be released in real-time through the data release method to provide strong monitoring data support for staff [9]. Currently, the traditional online monitoring data release method of composite submarine cables has a poor timeliness problem. When faced with a large amount of online monitoring data, it takes a long time to complete the data release task [10].

The motivation behind this work stems from the limitations of the conventional method of online composite submarine cable monitoring data release, which utilizes blockchain dynamic zoning consensus. This method has drawbacks such as omissions in data release and a lengthy process, leading to reduced timeliness of the released data. To overcome these limitations and improve the efficiency of data release, the authors propose a new data publishing method by introducing horizontal federation learning.

Horizontal federated learning can improve the above problems, and it has been a popular machine learning

Table 2.1: Technical performance parameter settings of BOTDA series data acquisition optical fiber sensors

No	Technical parameter	Technical indicators
1	Fiber type	Standard single mode fiber
2	Optical fiber configuration	Dual fiber loop
3	Measuring distance	5km-100km
4	Measurement time	As low as 20s
5	Temperature measurement range	-190 °C ~ 700 °C (depending on the fiber optic cable material)
6	Temperature measurement accuracy	1 °C (full range)
7	Temperature resolution	0.1 °C
8	Strain measurement range	-4000 $\mu$ E to 6000 $\mu$ E (depending on cable material)
9	Strain measurement accuracy	$\pm 20 \mu$ E (full journey)
10	Strain resolution	2 $\mu$ E
11	Sampling interval	0.1-1m

method for multi-device joint modeling in recent years. In horizontal federated learning, the data provider can train the machine model [21] on the device according to the task published by the data requester. Then, the data provider will update the local model instead of the original data to the central server for global aggregation. Based on this, this paper introduces horizontal federated learning and proposes a new online monitoring data publishing method.

The main research question relies on:

"How can the timeliness of online composite submarine cable monitoring data release be improved by introducing horizontal federation learning?"

Additionally, the following sub-questions can be explored:

1. How can the online monitoring data of composite submarine cables be collected and preprocessed to eliminate the high-frequency capacitive effect?
2. How can the management of composite submarine cable data nodes be optimized to transform the status relationship of data nodes and ensure data quality?
3. How can a horizontal federation learning model be established to design the online monitoring data release process?
4. What are the experimental results of the proposed data release method in terms of feasibility, time required for data release, and timeliness of the released data?

## 2. Design of data release method for online monitoring of composite submarine cable.

**2.1. Composite submarine cable online monitoring data acquisition.** In the online monitoring data release method of composite submarine cable designed in this paper, online monitoring data collection is essential. It is the fundamental guarantee to ensure the accuracy of subsequent data release. First, select the data acquisition equipment [25] that meets the requirements of online monitoring data acquisition of composite submarine cables. After comprehensively considering the performance and operation characteristics of the data acquisition equipment, this paper selects the BOTDA series data acquisition optical fiber sensor as the acquisition equipment. Its technical performance parameters are shown in Table 2.1.

As shown in Table 2.1, these are the technical performance parameters of the composite submarine cable online monitoring data acquisition sensor selected in this paper. The data acquisition equipment collects real-time composite cable online monitoring data [14] under various working conditions. Due to the existence of capacitance to ground in each phase of the power cable, the capacitive current on the line is not equal during operation, and there is a high-order harmonic current in the line; when the cable is laid for a long distance, the capacitance current will be more discrete [16]. Therefore, it is necessary to process the online monitoring data of submarine cables to eliminate the high-frequency capacitive effect of submarine cables. The coaxial capacitor calculation method is used to characterize the capacitance characteristics of each phase of submarine cables and then judge the elimination of high-frequency capacitive effect [13]. The calculation expression of

capacitance characteristics of each stage of submarine cable is:

$$C = \frac{2\pi\varepsilon_0\varepsilon_r}{\ln(R_2/R_1)}$$

Among them,  $\varepsilon_0$ ,  $\varepsilon_r$  represent the dielectric constant of submarine cable insulator respectively;  $R_1$ ,  $R_2$  represent the outer radius and inner radius of submarine cable insulator respectively. After the elimination of the high-frequency capacitive effect of the submarine cable, it automatically communicates with the BOTDA data acquisition sensor through Ethernet and acquires and stores the monitoring data such as the optical fiber temperature, strain information, and power harmonic information collected by the equipment [5].

In the online monitoring data release method of composite submarine cables proposed in this paper, the online monitoring data collection process is crucial for ensuring the accuracy of subsequent data release. This discussion will focus on the techniques used for online monitoring data collection, with specific emphasis on selecting appropriate data acquisition equipment.

To begin with, selecting the right data acquisition equipment is essential to meet the requirements of online monitoring data acquisition for composite submarine cables. The choice of equipment plays a vital role in accurately capturing and recording the necessary data. In this work, the authors consider various factors, such as performance and operational characteristics, before finalizing the equipment selection.

One specific equipment selected in this study is the BOTDA (Brillouin Optical Time-Domain Analysis) series data acquisition optical fiber sensor. The selection is based on a comprehensive evaluation of its technical performance parameters, provided in Table 2.1. These parameters likely include measurement range, resolution, sampling rate, accuracy, and sensitivity. The BOTDA series data acquisition optical fiber sensor is known for capturing precise and reliable data from composite submarine cables, making it suitable for the proposed online monitoring data collection process.

The selection of appropriate data acquisition equipment ensures that the collected data is high quality and accurately represents the monitoring parameters of interest. Using the BOTDA series data acquisition optical fiber sensor, the authors aim to gather reliable and relevant data from composite submarine cables, which will serve as the foundation for subsequent data release and analysis.

**2.2. Composite submarine cable data node management.** After completing the online monitoring data collection of composite submarine cables, next, manage the collected online monitoring data nodes, transform the status relationship of data nodes, and ensure the quality of subsequent data release. A data node is a device that communicates with the system manager to complete data transmission, providing a running and maintenance environment [11]for database agents and data applications. Data providers who want to use the system to manage and share data sources need to create data nodes and add them to the system cluster network [1]in addition to acquiring the role of the system's data provider. The data node management module provides data node registration network, status detection, node authentication token management, and disable enable functions [23]. As shown in Figure 2.1, it is the state transformation relationship of data nodes designed in this paper.

According to the status transformation relationship of online monitoring data nodes shown in Figure 2.1, each data node is comprehensively transformed. The data node management module at the system management end receives the node heartbeat from the node status detector running at the data node end, confirms that the data node has successfully entered the network and is in a healthy running state, updates the node status to running, and records when the token expires [12]. After that, the system data node management module regularly queries the node status through the node status detector and updates it in the data node table [2]. The authentication token of the data node is valid within one week after the node successfully enters the network. The system data node management module regularly updates the permits of all online data nodes and their validity (the node status is Running) [7]. When updating the ticket, the discarded node will be found (the node status is Down), and the token update and node validity will be skipped [19]. Suppose the node reconnects within the token validity period. In that case, the node status is updated to Running, and the node that has not reconnected beyond the token validity period is judged invalid (the node status is invalid), the data provider needs to re-operate the data node into the network [24]. On this basis, the data node table maintained by

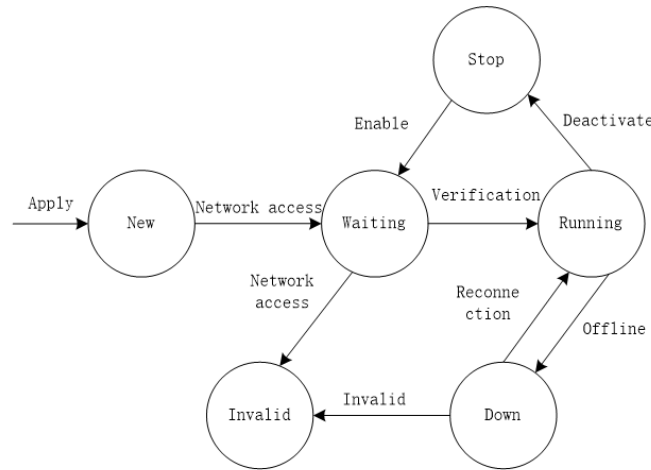


Fig. 2.1: Status transformation relationship of composite submarine cable online monitoring data nodes

Table 2.2: Maintenance data nodes of composite submarine cable online monitoring data node management module

Field Name	Field Type	Describe
Data Node ID	Int	Data node ID, a unique identifier of the node.
Data Node IP	String	Data node P address.
Name	String	Data node name.
User ID	Int	User ID. The data provider user ID to which the data node belongs.
Package Uri	String	The storage path of the data node access kit in Minio.
Token	String	Data node authentication token.
Toke Valid Time	Time	Token validity period.
Status	String	Data node status.

the composite submarine cable online monitoring data node management module is constructed, as shown in Table 2.2.

According to the maintenance node table structure shown in Table 2.2, after the data provider deactivates the data node, it updates the node status to "Stopped". It clears the node authentication token until the data provider enables the data node to reapply for the node authentication token and update the validity of the ticket, and the node status.

**2.3. Online monitoring data release based on horizontal federal learning.** After the management of the above composite submarine cable online monitoring data nodes is completed, on this basis, the research on online monitoring data release is carried out using the horizontal federation learning principle. Federal learning aims to solve the problem of joint knowledge for decentralized data sets. For all parties involved in federal training, "data is available but not visible to each other". First, a horizontal federation learning model is established, and the data processing of horizontal federation learning is obtained through the model structure. The flat federation learning model based in this paper is shown in Figure 2.2.

As shown in Figure 2.2, the horizontal federated learning model is applicable to the situation where the data characteristics of the data owned by each participant overlap. The data characteristics of each participant are aligned and the data samples are different, which can also be called federated learning by sample. In the training process of the horizontal federation model, the definition of system security is usually based on the



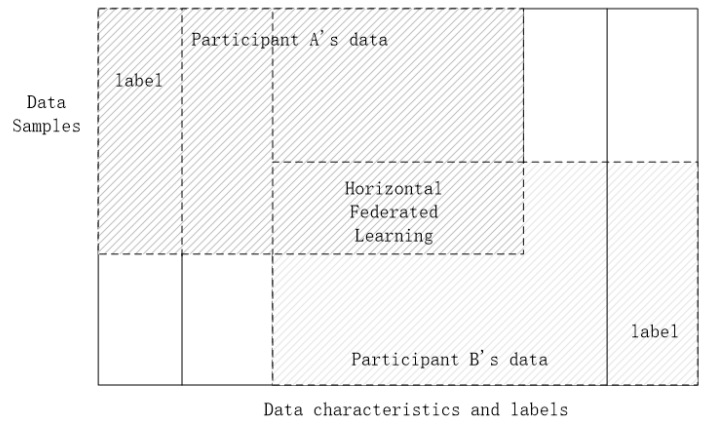


Fig. 2.2: Building a horizontal federation learning model

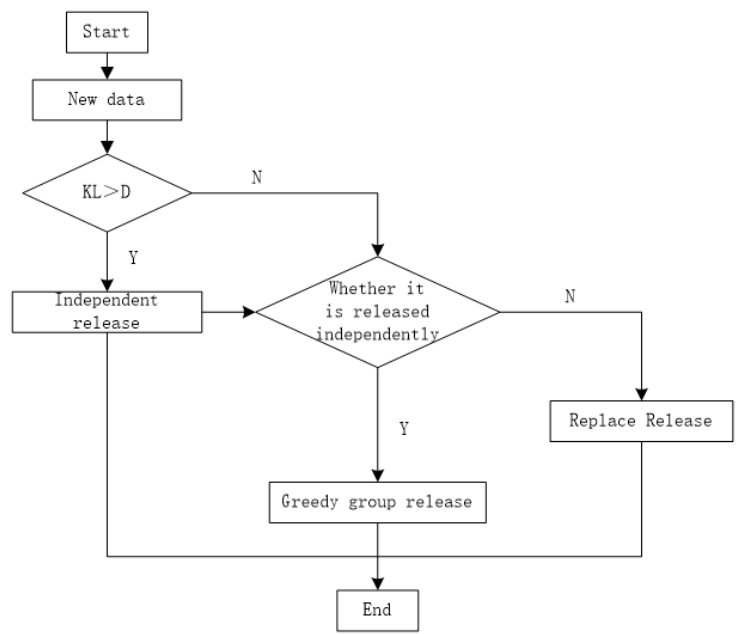


Fig. 2.3: Release process of composite submarine cable online monitoring data based on horizontal federated learning

assumption that the participants in the flat federation learning system are honest and trustworthy and pay attention to the Fusion Server [15] whose intermediate results converge. In order to protect the privacy and integrity of online monitoring data released by composite cables, it is necessary to avoid storing the eigenvalues of the model in clear text, and usually use secure multi-party computing and homomorphic encryption to ensure the security of intermediate results.

The online monitoring data release process of composite submarine cable based on horizontal federated learning designed in this paper is shown in Figure 2.3.

As shown in Figure 2.3, first use KL divergence to calculate the difference between the updated data at the current time point and the data published at the previous time point, and then add noise to the divergence value

Table 3.1: Description of Thermal Resistance Coefficient of Structural Component Materials of Each Layer of Submarine Cable

No	Submarine cable structure	Material thermal resistance coefficient/ $K \cdot m/W$
1	Conductor	0.00
2	Peninsula resistance water zone	6.00
3	Bituminous PP rope cover layer	6.00
4	Asphalt anticorrosive coating	5.00
5	Insulation shield	3.50
6	Lead alloy sheath	0.03
7	Conductor shield	3.50

and predefined threshold. Compare the noise KL divergence of online monitoring data of composite submarine cable with the noise threshold  $D$ , and then judge how the data at the current time point should be released. If the KL divergence of the online monitoring data noise is greater than the noise threshold, the data shall be removed by independent data release method; If the noise KL divergence of online monitoring data is less than the noise threshold, the second judgment process will be started to judge whether the online monitoring data at the last adjacent time point is released independently. If the data at the last adjacent time point is released separately, the data will be released in greedy groups if the data at the last adjacent time point is not released separately. Then this data release will adopt the alternative release method to complete the online monitoring data release task of composite submarine cables. When some online monitoring data nodes are merged, their data sensitivity will also be reduced to varying degrees. After the above judgment and the selection of data release mode, the data can be released in an all-around way while meeting the timeliness and accuracy of data release.

### 3. Experimental analysis.

**3.1. Experiment preparation.** In order to test the feasibility of the online monitoring data publishing method of composite submarine cable based on horizontal federal learning proposed in this paper, and ensure its publishing effect, the experimental analysis is carried out as shown below. Firstly, according to the mechanical test of submarine cable, the test environment is arranged, and the winding test device, linear tensile test device and tensile bending test device are selected respectively. In the winding test device, the inner diameters of the two submarine cable drums are set as 5m and 6m respectively, and the spacing between the two submarine cable drums is set as 15.8m. The submarine cable rises from one cable barrel and enters the other through the fixed pulley above the other cable pillar and the fixed pulley above the other cable barrel. The height of the fixed pulley is 14m, and the distance between two fixed pulleys is 21.2m. The maximum tension of the direct pulling device and the tension bending device is 35kN, and the diameter of the tension bending device drum is 6.5m. The submarine cable adopts YPLKY42110KV single-core optical fiber composite submarine cable. The length of the winding test cable is 220 meters, and the length of the straight tension and tension bending test cable is 27.6 meters. The tension of the submarine cable is measured by a tensiometer, and the strain and temperature of the composite optical fiber in the submarine cable are measured by BOTDR and ROTDR.

Due to the complex structure of submarine cable, if it is divided into four layers: conductor layer, insulation layer, inner lining layer, and outer protective layer, it will inevitably lead to large calculation errors. Only the layers with the same material thermal resistance coefficient are combined to take into account the simplification of the thermal path and calculation accuracy. In this experiment, the thermal resistance coefficient of structural component materials of each layer of submarine cable is shown in Table 3.1.

As shown in Table 3.1, it is the thermal resistance coefficient of the structural component materials of each layer of the submarine cable. On this basis, the online monitoring data release method of composite submarine cable based on horizontal federated learning proposed in this paper is used to release the data and test the application effect of the release method.

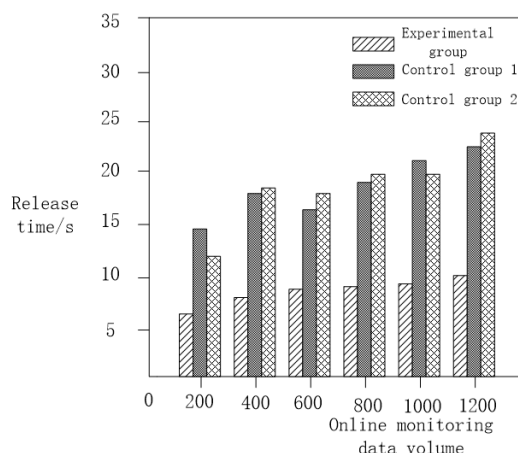


Fig. 3.1: Comparison Results of Online Monitoring Data Release Time of Composite Submarine Cable

**3.2. Result analysis.** In order to make the test results of this experiment have corresponding persuasiveness, this paper specifically sets up two control groups, of which control group 1 is the monitoring data publishing method based on differential privacy, control group 2 is the monitoring data publishing method based on distributed optical fiber strain, and the monitoring data publishing method based on horizontal federated learning proposed in this paper is the experimental group. Through comparison, the feasibility of the proposed data release method is verified. The time required for the online monitoring data of composite submarine cables to complete the release task is selected as the evaluation index of this experiment. The shorter the time required for data release, the higher the feasibility of the data release method, and the better the timeliness of the release. Set the amount of online monitoring data of composite submarine cable as 200, 400, 600, 800, 1000, and 1200, use MATLAB simulation analysis software to simulate the release process of the three data release methods, using SPSS statistical analysis software to measure and integrate the time required for the three methods to complete the data release task, and draw a comparison chart of evaluation indicators.

It can be seen from the comparison results of the evaluation indicators in Figure 3.1 that the time required for the three data release methods to complete the online monitoring data release task is quite different when the amount of online monitoring data of composite submarine cables is gradually increasing. Among them, the data publishing method based on horizontal federal learning proposed in this paper shows good advantages. In each group of online monitoring data publishing tasks of composite submarine cables, the time required to complete the publishing is less than the other two methods, and the publishing timeliness is good, without a significant growth trend. While control group 1 and control group 2 took longer to complete the release task with the increase of online monitoring data. From the comparison results, it can be seen that the data release method proposed in this paper is highly feasible and more suitable for the online monitoring data release of composite submarine cables.

**3.3. Result discussion.** In this context, the data publishing method based on horizontal federation learning, as proposed in this paper, demonstrates clear advantages. Specifically, this method outperforms the other two methods in terms of completion time and timeliness of data release. For each group of online monitoring data publishing tasks for composite submarine cables, the proposed method consistently requires less time to complete the data release compared to the control groups (control group 1 and control group 2). Furthermore, the timeliness of the data release using the proposed method remains good and does not show a significant growth trend. In contrast, the control groups experience longer completion times as the amount of online monitoring data increases.

Based on the comparison results, it is evident that the data release method proposed in this paper is highly feasible and well-suited for online monitoring data release of composite submarine cables. The advantages of

the proposed method, such as faster completion times and consistent timeliness, highlight its effectiveness in addressing the challenges associated with large volumes of online monitoring data.

These results imply that the proposed method based on horizontal federation learning offers a practical solution for efficient and timely data release in the context of composite submarine cable monitoring. Its superiority over the control groups suggests that it can effectively handle the increasing amount of online monitoring data, making it a favorable choice for composite submarine cable data release tasks.

**4. Conclusion.** The release of online monitoring data of composite submarine cables is of great significance for ensuring power transmission between coastal islands and cities. In order to improve the poor timeliness of traditional data release methods and the inability to release online monitoring data of composite submarine cables in time, this paper proposes a new method of online monitoring data release of composite submarine cables by introducing horizontal federation learning based on the traditional data release methods. Through the research in this paper, the time required for data release is reduced, the timeliness of data release is improved, and the monitoring data released can timely identify potential fault hazards in the operation of submarine cables, reduce the probability of damage to submarine cables, and ensure the safety of power grid operation.

#### REFERENCES

- [1] M. ASOODAR, M. NAHALPARVARI, C. DANIELSSON, R. SÖDERSTRÖM, AND H.-P. NEE, *Online health monitoring of dc-link capacitors in modular multilevel converters for facts and hvdc applications*, IEEE Transactions on Power Electronics, 36 (2021), pp. 13489–13503.
- [2] C. BREHL, H. U. BRASS, C. LÜCHTRATH, L. BÖCKMANN, N. IHLING, T. CLASSEN, J. PIETRUSZKA, AND J. BÜCHS, *Optimized prodigiosin production with pseudomonas putida kt2440 using parallelized noninvasive online monitoring*, Biotechnology Progress, 38 (2022), p. e3245.
- [3] Y. CHENG, *Monitoring method based on gis for submarine cable fault data (retraction of vol 14, art no 490, 2021)*, Arabian Journal of Geosciences, (2021).
- [4] H. DELVA, J. A. ROJAS, P. COLPAERT, AND R. VERBORGH, *Geospatially partitioning public transit networks for open data publishing*, Journal of Web Engineering, (2021), pp. 1003–1026.
- [5] X. DUAN, J. ZOU, B. LI, Z. WU, AND D. LEI, *An online monitoring scheme of output capacitor's equivalent series resistance for buck converters without current sensors*, IEEE Transactions on Industrial Electronics, 68 (2020), pp. 10107–10117.
- [6] B. M. T. FOUDA, B. YANG, D. HAN, AND B. AN, *Pattern recognition of optical fiber vibration signal of the submarine cable for its safety*, IEEE Sensors Journal, 21 (2020), pp. 6510–6519.
- [7] H. HONG, Z. CAI, W. WANG, AND Y. LIU, *An online monitoring method for creep-fatigue life consumption with real-time damage accumulation*, International Journal of Damage Mechanics, 30 (2021), pp. 764–785.
- [8] M. S. KHAN, A. ANJUM, T. SABA, A. REHMAN, AND U. TARIQ, *Improved generalization for secure personal data publishing using deviation*, IT Professional, 23 (2021), pp. 75–80.
- [9] U. KÜPPER, T. HERRIG, A. KLINK, D. WELLING, AND T. BERGS, *Evaluation of the process performance in wire edm based on an online process monitoring system*, Procedia CIRP, 95 (2020), pp. 360–365.
- [10] S. LI, C. TONG, Y. CHEN, AND T. LAN, *Dynamic statistical process monitoring based on online dynamic discriminative feature analysis*, Journal of Process Control, 103 (2021), pp. 67–75.
- [11] T. LI, J. CHEN, P. CONG, X. DAI, R. QIU, AND Z. LIU, *Online condition monitoring of dc-link capacitor for ac/dc/ac pwm converter*, IEEE Transactions on Power Electronics, 37 (2021), pp. 865–878.
- [12] Y. LI, Y. LIU, Y. TIAN, Y. WANG, AND J. WANG, *Application of improved fireworks algorithm in grinding surface roughness online monitoring*, Journal of Manufacturing Processes, 74 (2022), pp. 400–412.
- [13] P. A. LIRA-PARADA, A. TUVERI, G. M. SEIBOLD, AND N. BAR, *Comparison of noninvasive, in-situ and external monitoring of microbial growth in fed-batch cultivations in corynebacterium glutamicum*, Biochemical Engineering Journal, 170 (2021), p. 107989.
- [14] A. LIU, J. YI, X. DING, J. DENG, D. WU, Y. HUO, J. JIANG, Q. LI, AND J. CHEN, *An online technology for effectively monitoring inorganic condensable particulate matter emitted from industrial plants*, Journal of Hazardous Materials, 428 (2022), p. 128221.
- [15] F. LIU AND Y. TANG, *Research on data transmission of wireless sensor network based on feedback cooperation*, Computer Simulation, (2022), p. 039.
- [16] G. LU, D. ZHENG, Q. ZHANG, AND P. ZHANG, *Effects of converter harmonic voltages on transformer insulation ageing and an online monitoring method for interlayer insulation*, IEEE Transactions on Power Electronics, 37 (2021), pp. 3504–3514.
- [17] K. MCGLINN, R. BRENNAN, C. DEBRUYNE, A. MEEHAN, L. MCNERNEY, E. CLINTON, P. KELLY, AND D. O'SULLIVAN, *Publishing authoritative geospatial data to support interlinking of building information models*, Automation in Construction, 124 (2021), p. 103534.
- [18] M. MUNEEZ, *Protection of submarine optical fibre cables in the strait of malacca.*, Journal of Marine Environmental Engineering, 10 (2021).
- [19] K. V. RAO, Y. P. KUMAR, V. K. SINGH, AND B. S. PRASAD, *Online modeling and monitoring of power consumption, aerosol emissions, and surface roughness in wire cut electric discharge machining of ti-6al-4 v*, The International Journal of

- Advanced Manufacturing Technology, (2022), pp. 1–18.
- [20] Q. REN, M. LI, T. KONG, AND J. MA, *Multi-sensor real-time monitoring of dam behavior using self-adaptive online sequential learning*, Automation in Construction, 140 (2022), p. 104365.
  - [21] X. WANG, Z. WANG, E. TAHCHI, AND M. ZUKERMAN, *Submarine cable path planning based on weight selection of design considerations*, IEEE Access, 9 (2021), pp. 123847–123860.
  - [22] Y. WU AND P. ZHANG, *Online monitoring for power cables in dfig-based wind farms using high-frequency resonance analysis*, IEEE Transactions on Sustainable Energy, 13 (2021), pp. 378–390.
  - [23] J. XIN, M. DU, Z. OUYANG, AND K. WEI, *Online monitoring for threshold voltage of sic mosfet considering the coupling impact on bti and junction temperature*, IEEE Transactions on Electron Devices, 68 (2021), pp. 1772–1777.
  - [24] X. ZHU, D. ZHANG, AND W. GAO, *Online noninvasive technique for condition monitoring of capacitor in linear power supplies*, IEEE Transactions on Power Electronics, 36 (2021), pp. 8761–8773.
  - [25] S. ZHUANG, Q. WU, Y. WANG, Y. MA, L. LI, G. ZHANG, Y. CHEN, Y. ZHANG, H. LIU, X. QI, ET AL., *Online monitoring of air activation at the china spallation neutron source*, Journal of Environmental Radioactivity, 234 (2021), p. 106642.

*Edited by:* Sathishkumar V E

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

*Received:* May 13, 2023

*Accepted:* Jul 10, 2023





## REAL-TIME MONITORING METHOD OF INSULATION STATUS OF PHOTOELECTRIC COMPOSITE SUBMARINE CABLE BASED ON THERMOELECTRIC COUPLING

**Abstract.** The existing approach for monitoring the insulation state of photoelectric composite submarine cables primarily relies on detecting the current of the cable protection layer. However, this conventional method suffers from limited monitoring accuracy due to the absence of parameter identification processing for the cable. As a result, there is a need to improve the monitoring methodology by incorporating robust parameter identification techniques to enhance the accuracy of insulation state evaluation. In this regard, a real-time monitoring method based on thermoelectric coupling is proposed to monitor the insulation status of the photoelectric composite submarine cable. By constructing an equivalent composite circuit model and a thermodynamic function, a thermoelectric coupling model is constructed and used to identify the parameters of the submarine cable; by extracting the frequency extremes in the spectral values of the submarine cable current signal, an equivalent insulation characteristic function is constructed to realize the determination of the insulation state. The proposed method is verified for the insulation state monitoring effect in the experiment. The experimental results show that when the proposed method is used to monitor the insulation state of the photoelectric composite submarine cable, the calculated partial discharge quantity has a small error, and the monitoring accuracy is high.

**Key words:** thermoelectric coupling model; equivalent circuit; photoelectric composite submarine cable; insulation condition monitoring;

**1. Introduction.** there are three main methods for monitoring the insulation state of photoelectric composite submarine cable: partial discharge monitoring method, insulation protection layer monitoring method and cable temperature monitoring method. Among them, the partial discharge monitoring method is mainly through the sensing equipment of the discharge signal issued by the photoelectric composite submarine cable to collect and combined with filtering algorithms for the discharge signal discrete processing to achieve signal denoising based on identifying the frequency peak in the discharge signal curve, determine the number of partial discharge of the photoelectric composite submarine cable in the sampling period [1]. At the same time, combined with the change of frequency amplitude of the discharge signal, it can also be roughly inferred from the discharge signal issued by the specific node coordinates. The insulation protection layer monitoring method is mainly through the power outage state of the protection layer in the photoelectric composite submarine cable residual current value detection. The construction of the equivalent current model will be compared with the normal form of the current if the present error is large, it means that at this moment, the photoelectric composite submarine cable insulation aging situation needs to repair and replace the treatment [15]. The cable temperature monitoring method is mainly through the infrared imaging method to capture the heat of the unit length of the photoelectric composite submarine cable and the formation of infrared images through the cable in the abnormal temperature fluctuations node to determine the cable insulation aging state. The above three methods can be achieved to a certain extent to monitor the insulation status of the photoelectric composite submarine cable, but there are certain limitations. First, the temperature monitoring method can be applied to a small range, only to the middle part of the cable, for the cable joints, and the temperature of the connection with other electrical equipment can not achieve more accurate monitoring. This research focuses on finding strong photoelectric composite submarine cable insulation material and studying its efficiency.

The insulation layer protection monitoring method needs to be carried out under the scene of power failure, which is not conducive to the regular work of the photoelectric composite submarine cable and quickly affects production efficiency. This method can only determine whether the insulation aging situation, the specific breakdown of the insulation state can not be accurately identified, and can not get the particular insulation node coordinates, so it is also not conducive to the maintenance of photoelectric composite submarine cable [14]. The partial discharge monitoring method not only can identify the different states of insulation but also can get the specific partial discharge node coordinates by combining the fluctuation of the discharge amount

to achieve the identification of the insulation point location, which has a more excellent monitoring effect. Therefore, this paper combines the conventional partial discharge monitoring method through the optimization of the traditional process and the introduction of the thermoelectric coupling model to identify the parameters of the photoelectric composite submarine cable. For the interference of the conventional way in terms of noise, a filtering algorithm is used to discretize the noise of the discharge signal and improve the processing effect of the discharge signal [17]. The major research challenge is as follows:

1. Submarine cables must be designed to accommodate high transmission capacities to meet the growing demand for data and communication services across continents and undersea regions.
2. Submarine cables are vulnerable to physical damage caused by fishing activities, anchoring, natural disasters, and geological hazards such as earthquakes and landslides. Designing adequate protection measures, such as armoring, shielding, and burial, is essential to safeguard the cables against these potential risks.
3. Submarine cables can span hundreds or even thousands of kilometers across various terrains, including deep ocean trenches and shallow coastal areas. Designing cables that can withstand the stresses and strains associated with installation, deployment, and maintenance in diverse environments is crucial.
4. Submarine cables face signal degradation due to attenuation, dispersion, and noise. Design considerations include minimizing signal loss, optimizing signal integrity, and employing repeaters and amplifiers to maintain signal quality over long distances.
5. Submarine cable installations can impact marine ecosystems and habitats. Mitigating environmental impacts through careful route selection, installation techniques, and adherence to environmental regulations is an important aspect of cable design.
6. Submarine cables are critical infrastructure for global communication networks. Ensuring high reliability and redundancy is crucial to minimize downtime and service disruptions. Designing robust cable systems with redundant paths, fault detection and localization capabilities, and quick restoration mechanisms is essential.
7. Accessing and repairing submarine cables in deep-sea environments pose significant logistical and technical challenges. Designing cables that facilitate efficient maintenance and repair processes, including remotely operated vehicles (ROVs) and repair ships, is essential to minimize downtime and service disruptions.
8. Submarine cable projects involve substantial manufacturing, installation, and maintenance investments. Balancing the cost with performance, lifespan, and other requirements is a key challenge in designing economically viable submarine cable systems.

The main contribution of the research is performing real-time monitoring of thermo electric coupling of the submarine cable with insulating solid advantage.

**2. Photoelectric composite submarine cable parameter identification based on thermoelectric coupling model.** The photoelectric composite submarine cable is affected by different electromagnetic fields in the working state, so it has more complex physical characteristics. To accurately capture the insulation state of the photoelectric composite submarine cable, it is necessary to construct its temperature function first. In this regard, this paper combines the thermoelectric coupling model to identify the parameters of the photoelectric composite submarine cable. It provides help for the subsequent online monitoring of the insulation state [2]. Therefore, when monitoring the insulation state of photoelectric composite submarine cable, not only the circuit state can be monitored, but also the heat change can be observed by the temperature function, which is combined with the equivalent circuit model and thermodynamic model to build the thermoelectric coupling model shown in Figure 2.1 [7].

The optical composite submarine cable will generate and transfer heat in the working state, and in the insulated state, the heat will also appear transient, so combined with the law of conservation of energy, this paper describes the heat transfer process of the optical composite submarine cable as follows [16].

$$\rho_i C_i \frac{\sigma T}{t} = \nabla \cdot (\lambda_i \nabla T) + q \quad (2.1)$$

In the above equation, the heat gain of a submarine cable in the insulated state is equal to the product of the rate of heat production of the wire inside the line and the combined heat generated when the submarine



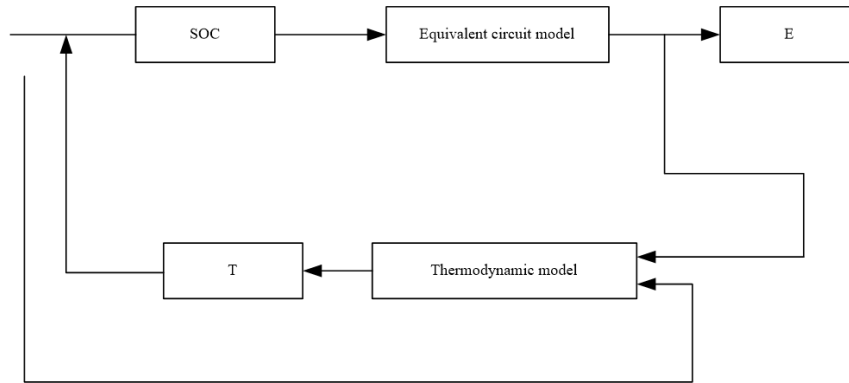


Fig. 2.1: Thermoelectric coupling framework

Table 2.1: Equivalent circuit model identification parameters

Model parameters	Expressions
Open circuit voltage	$R_o$
Open circuit current	$R_e$
SOC	$R_d$
Submarine cable resistance	$C_e$
Submarine cable capacitance	$C_d$

cable wire interacts with the outside world. Where,  $\rho_i$  represents the density of the submarine cable wires,  $C_i$  represents the specific heat capacity of the insulation,  $T$  represents the average temperature of the wires,  $q$  represents the thermal conductivity of the whole line,  $\lambda_i$  represents the rate of heat generation of the submarine cable, and  $\sigma$  represents the number of the submarine cable wires. For the above thermoelectric coupling model, the conventional CFD simulation software can be used to solve and analyze it. Still, because the software cannot effectively identify the multi-parameter mathematical model, it takes a long time to analyze using the software, which quickly affects the subsequent monitoring efficiency [9]. There are two ways to optimize the model: one is to change the order of the model and reduce the demand of the model, thus reducing the analysis time of the software; the other is to decompose the transient conditions of the model and divide it into different stages, which are input into CFD software for solving and analyzing. Since the second method is more complicated to implement, this paper chooses the first method to reduce the order of the thermoelectric coupling model to achieve effective non-capture of the operating state of the photoelectric composite submarine cable [10]. The literature shows that this method has the same data processing effect as the CFD model, and it takes less time to complete the simulation analysis of the thermoelectric coupling model in a limited time, which can improve the subsequent insulation state monitoring efficiency. The parameters to be identified and their corresponding mathematical expressions are shown below in the equivalent circuit model of the photoelectric composite submarine cable.

In order to improve the accuracy of the parameter identification of the equivalent circuit model, this paper chooses to combine the least squares method to estimate the parameters in the above table and then combine them with the actual observations to achieve the online identification of the parameters [11]. Since the parameters of the equivalent circuit model need to be reformatted for each set of data when using the conventional method, which may affect the efficiency of the model, this paper adopts a recursive method to process them, and the specific processing formula is shown below.

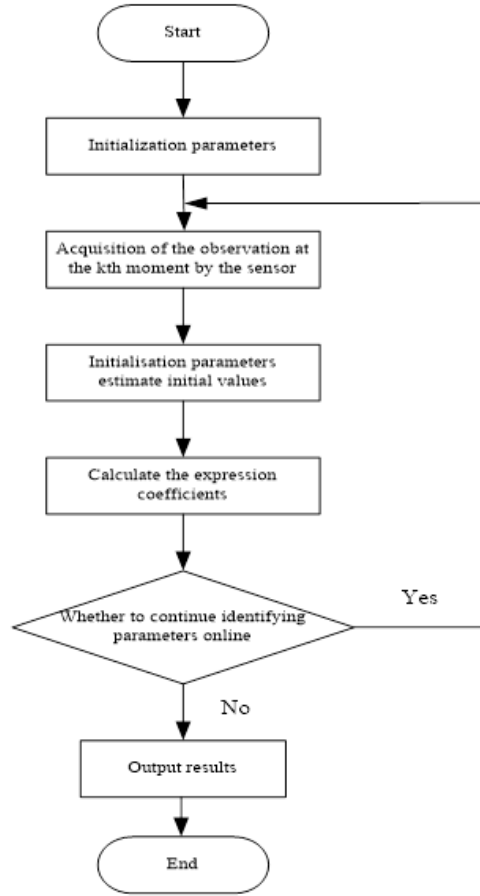


Fig. 2.2: Flow chart of parameter identification

$$\begin{cases} \theta(k) = \theta(k-1) + K(k)[y(k) - \varphi(k)] \\ K(k) = P(k)\varphi(k) \\ P(k) = [1 - K(k)\varphi(k)]P(k-1) \end{cases} \quad (2.2)$$

Where,  $\theta(k)$  represents the predicted value of the least squares method for the parameters of the equivalent circuit model,  $y(k)$  represents the actual observed value of the model,  $\varphi(k)$  represents the output and input values of the equivalent circuit,  $\varphi^T(k)$  represents the circuit data format conversion matrix,  $K(k)$  represents the recursive gain parameters, and  $P(k)$  represents the root mean square error matrix of the recursive results [13]. The above principle is applied to the thermoelectric coupling model constructed in this paper, assuming that the input value is the internal equivalent current flow value of the photoelectric composite submarine cable, representing the output value, and the frequency domain value of the model can be obtained by discretizing its transformation, and the specific formula is shown below.

$$y(k) = \alpha y(k-1) + \beta y(k-2) \quad (2.3)$$

The above equation can be used to identify the parameters of the equivalent circuit model, and the actual model parameters can be predicted by combining the recursive results with the actual observed values.

As can be seen from the above figure, the model parameter identification first requires initialization of the parameters, including the data sampling period of the photoelectric composite submarine cable, the open-circuit

Table 3.1: Optical composite submarine cable insulation state division results

Insulation status	Relative dielectric constant
intact	2.46
Slightly aged	2.57
Moderate ageing	2.63
Severe ageing	2.69
Damaged insulation	2.78

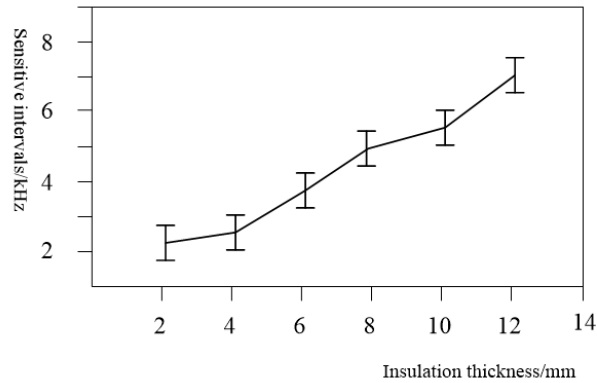


Fig. 3.1: Variable parameters on submarine cable sensitive frequency

voltage of the cable, and the state of charge [8]. Then, the  $k$ th moment observation data, including voltage data and current data, are obtained through the sensor, and the equivalent parameters are estimated in real-time. The open-circuit voltage of the optical composite cable at time  $k$  is calculated by the correspondence between the open-circuit voltage and the equivalent parameters. Then the recursive least squares method is enabled to perform recursive operations on the parameters of the model and calculate the expression parameters to obtain the online identification results at [6]. Through the above steps, the thermoelectric coupling model of the photoelectric composite submarine cable can be constructed, and the recursive least squares method can be used to identify the parameters of the model and provide data for the subsequent online identification of the insulation state of the submarine cable [18].

### 3. Optical composite submarine cable insulation eigen function extreme frequency extraction.

To effectively monitor the insulation state of the photoelectric composite submarine cable, this paper takes the common AC submarine cable in the power system as an example and divides the relative dielectric constant under its insulation state, by using the relative dielectric constant under different values to judge the insulation state of the submarine cable, the specific division results are shown in Table 3.1 [5].

Through the above division results can be seen when the relative dielectric constant of the photoelectric composite submarine cable is 2.78, which means that at this moment the submarine cable appears insulation damage situation. In this regard, this paper analyzes the variable parameters of the submarine cable to explore the influence of different variable parameters on the sensitive frequency of the cable, the specific analysis results are shown in Figure 3.1 [3].

Through the study of the relevant literature, it is known that the length change of the cable is negatively correlated with its sensitive frequency, while the radius of the cable core and the thickness of the insulation wrapping layer are positively correlated with the sensitive frequency of the cable. For an installed photoelectric composite submarine cable, its sensitive frequency variation interval is a definite fixed value, which can be solved by calculating the maximum and minimum values of the frequency variation curve. [4]. In this paper, we choose

Table 3.2: Reference values of extreme frequencies under different aging conditions

Insulation status	Relative dielectric constant
intact	7.05
Slightly aged	7.35
Moderate ageing	7.21
Severe ageing	7.16
Damaged insulation	6.75

to extract the frequency of the extreme value of the insulation characteristic function, combined with the least squares method, to analyze the variable parameters of the submarine cable, the specific implementation process is as follows.

In order to optimize the above process, this paper extracts the insulation spectrum by combining the high-frequency excitation generated outside the photoelectric composite submarine cable, by measuring the current flowing in the core inside the cable, by using the output value of the equivalent circuit model as the ground current, and by using the least squares method. According to the equivalent insulation function constructed above, this paper combines the cable parameters of the photoelectric composite submarine cable and inputs the parameters into the equivalent insulation function to obtain the reference values of the extreme frequency of the photoelectric composite submarine cable under different aging states, and the specific reference values are shown in Table 3.2.

According to the principle of equivalent insulation function, photoelectric composite submarine cable in different units of the length of the grounding resistance will not have an impact on the polar frequency, so this paper will be set to the value of the resistance of the core of the photoelectric composite submarine cable to 0, you can get for the polar frequency has a real impact on the equivalent insulation function, the specific function expression is shown below.

$$G_{in} = \lim G_0 = \left| \frac{y(z_c + z_s)}{z_c^2 + z_s^2} \right| \quad (3.1)$$

The above equivalent insulation function is consistent with the actual insulation state frequency extrema, so the actual insulation frequency extrema can be obtained by solving the above equivalent insulation function extrema [12].

**4. Optical composite submarine cable insulation state identification.** Optical composite submarine cable in the insulation state due to local aging and damage part of the resistance capacity is missing, so there will be a partial discharge situation, by capturing the partial discharge frequency, you can effectively determine the current insulation state of the optical composite submarine cable. This paper combines the above constructed insulation function and insulation frequency extremes to achieve effective monitoring of the insulation state by using the local discharge signal identification of the photoelectric composite submarine cable as the main means [19]. In general, the current pulse sensor can be used to detect the discharge signal generated by the local wear of the photoelectric composite submarine cable, but the local discharge signal detected by this method usually contains a complex signal including the external environment, the signal noise is large, so first need to remove the noise of the local discharge signal to obtain a pure discharge signal, in order to achieve the effective insulation state of the photoelectric composite submarine cable Identification.

Firstly, the current sensor is used to collect the discharge signal from the current flowing inside the photoelectric composite submarine cable; then the signal conditioning circuit is used to process the collected discharge signal, amplify the signal wave value, and use the denoising algorithm to discretize the collected signal to obtain the denoised signal. If the result shows that the extreme value of the discharge exceeds the set threshold, it means that the cable is aging and wearing insulation, and an alarm is issued to remind the staff to replace the insulated cable as soon as possible.

Through the above steps, the effective identification and early warning of the insulation state of the photoelectric composite submarine cable can be completed. Combined with the thermoelectric coupling model

Table 5.1: Optical composite submarine cable parameters

Parameter type	Voltage Rating			
	10kV	35kV	110kV	220kv
Wire core radius	8.2	5.5	8.6	8.9
Insulation thickness	4.5	10.5	12	16.5
Cable length	0-2000			

and the extreme frequency extraction of the insulation function constructed in this paper, the design of the real-time monitoring method of the insulation state of the photoelectric composite submarine cable based on thermoelectric coupling is now completed.

## 5. Experiment and analysis.

**5.1. Experimental preparation.** In order to prove that this paper proposes a real-time monitoring method based on the thermoelectric coupling of photoelectric composite submarine cable insulation state in the actual monitoring effect than the conventional photoelectric composite submarine cable insulation state real-time monitoring method, after the theoretical part of the design is completed, the experimental part is constructed to test the actual monitoring effect of this paper's method. In order to ensure the experimental effect, two conventional photoelectric composite submarine cable insulation state real-time monitoring methods are selected for comparison, namely, the photoelectric composite submarine cable insulation state real-time monitoring method based on electrical characteristics analysis, and the photoelectric composite submarine cable insulation state real-time monitoring method based on power disturbance.

The experimental object selected for this experiment is an offshore electrical engineering photoelectric composite submarine cable combination, the cable combination for different voltage levels has corresponding cable changeable parameters, and specific parameters configuration as shown in Table 5.1.

The above parameters of photoelectric composite submarine cable are used as input values, combined with MATLAB simulation software to construct a mathematical model of the submarine cable, and simulate different insulation states of the photoelectric composite submarine cable. Three methods are used to monitor the mathematical model in real-time and to compare the accuracy of the methods by comparing the error between the partial discharge amount and the actual amount under different methods.

**5.2. Analysis of test results.** The comparison standard selected for this experiment is the monitoring accuracy of the monitoring method, The specific measurement index is the error between the calculated partial discharge amount about the actual value under different methods, the smaller the error represents the higher the monitoring accuracy of the method, the more accurate monitoring of the insulation state of the photoelectric composite submarine cable, the specific experimental results are shown below.

The above experimental results show that the errors between the calculated partial discharge amount and the actual value under different partial discharge nodes and different methods are different. By observing the change curve of the partial discharge error, it is obvious that the partial discharge amount calculated by the method in this paper is closer to the actual value, while the error of the partial discharge amount under the conventional method is higher. Thus, it can be proved that the insulation condition monitoring method proposed in this paper has better monitoring accuracy.

**6. Conclusion.** This research paper addresses the issue of low accuracy in insulation state monitoring of conventional photoelectric composite submarine cables. By incorporating the thermoelectric coupling model and analyzing the thermodynamic and power transmission processes, the proposed approach effectively adjusts the insulation state function using error tolerance. This adjustment leads to calculated partial discharge quantities that closely align with actual results. By determining the magnitude and location of the partial discharge, the insulation state of the photoelectric composite submarine cable can be accurately identified and monitored.

However, it is important to acknowledge the limitations of this study. The research primarily relies on mathematical modeling and simulation. In particular, future work should concentrate on investigating the sig-

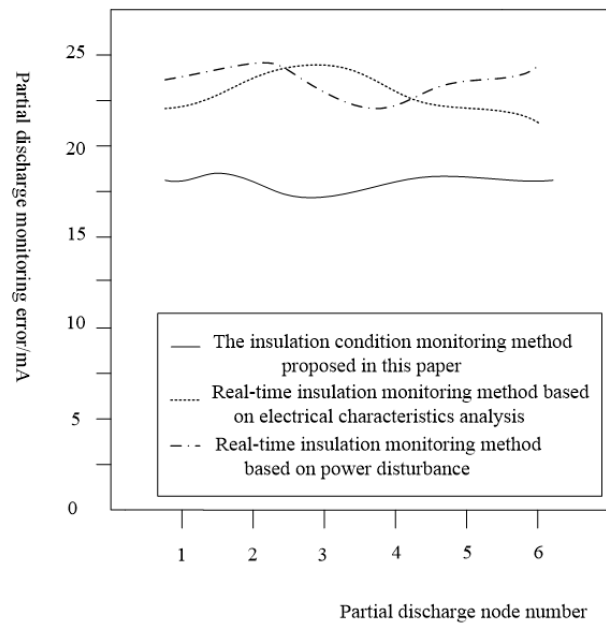


Fig. 5.1: Comparison of local discharge monitoring errors

nal's anti-interference capabilities to enhance the analysis of partial discharge signals. This research direction will contribute to improving the practical effectiveness of analyzing partial discharge signals and provide valuable assistance in maintaining photoelectric composite submarine cables. By addressing these limitations and conducting more extensive experimental validations, the proposed method can be refined and applied effectively in real-world submarine cable monitoring systems.

#### REFERENCES

- [1] H. CAO, Y. HAO, Z. ZHANG, J. WEI, AND L. YANG, *System and method of quasi-distributed fiber bragg gratings monitoring brittle fracture process of composite insulators*, IEEE Transactions on Instrumentation and Measurement, 70 (2021), pp. 1–10.
- [2] D. DU, N. GUO, L. ZHANG, Y. WU, Q. SHANG, AND W. LIU, *Real-time fluorometric monitoring of monophenolase activity using a matrix-matched calibration curve*, Analytical and Bioanalytical Chemistry, 413 (2021), pp. 635–647.
- [3] L. FAN, K. PENG, Z. ZHOU, Y. YAN, C. RAN, H. WANG, G. HAN, B. ZHANG, X. LU, G. WANG, ET AL., *Exceptional thermoelectric performance enabled by high carrier mobility and intrinsically low lattice thermal conductivity in phosphide  $cd_3p_2$* , Chemistry of Materials, 34 (2022), pp. 1620–1626.
- [4] S. H. HEO, J. YOO, H. LEE, H. JANG, S. JO, J. CHO, S. BAEK, S. E. YANG, D. H. GU, H. J. MUN, ET AL., *Solution-processed hole-doped  $sns_e$  thermoelectric thin-film devices for low-temperature power generation*, ACS Energy Letters, 7 (2022), pp. 2092–2101.
- [5] Y. JIN, R. TAO, S. LUO, AND Z. LI, *Size-sensitive thermoelectric properties of electrolyte-based nanofluidic systems*, The Journal of Physical Chemistry Letters, 12 (2021), pp. 1144–1149.
- [6] J. KULLIK, *Justin sherman: Cyber defense across the ocean floor. the geopolitics of submarine cable security. washington, dc: Atlantic council, september 2021, 2022.*
- [7] P. R. KUMAR, G. B. MOHAMMAD, AND P. DILEEP, *Real-time heart rate monitoring system using least square method*, Annals of the Romanian Society for Cell Biology, 25 (2021), pp. 16302–16308.
- [8] H. LI, P. YU, S. LI, AND J. DENG, *Study on internal overvoltages of 220 kv submarine cable transmission system of the offshore oil-gas field group*, Energy Reports, 8 (2022), pp. 377–386.
- [9] F. LIAO, Y. SHI, Q. DANG, H. YANG, H. HUANG, Z. KANG, AND M. SHAO, *Carbon dots dominated photoelectric surface in titanium dioxide nanotube/nitrogen-doped carbon dot/gold nanocomposites for improved photoelectrochemical water splitting*, Journal of Colloid and Interface Science, 606 (2022), pp. 1274–1283.
- [10] W. LIAO, W. WANG, D. SUN, Q. CUI, X. ZUO, Q. YANG, H. TANG, S. JIN, AND G. LI, *Optimized photoelectric catalysis with enhanced durability in rgo-interconnected nanocarbon-confined cobalt nanoparticles*, Carbon, 189 (2022), pp. 561–570.

- [11] M. MUNEEZ, *Protection of submarine optical fibre cables in the strait of malacca.*, Journal of Marine Environmental Engineering, 10 (2021).
- [12] W. NIU, X. CAO, Y. HU, F. WANG, AND J. SHI, *Approximate analytical solution to the temperature field in annular thermoelectric generator made of temperature-dependent material*, IEEE Transactions on Electron Devices, 68 (2021), pp. 6386–6392.
- [13] K. TAWESINTANANON, M. LANDRØ, J. K. BRENNE, AND A. HAUKANES, *Distributed acoustic sensing for near-surface imaging using submarine telecommunication cable: A case study in the trondheimsfjord, norway*, Geophysics, 86 (2021), pp. B303–B320.
- [14] M. M. VISSER, S. CHARLEER, S. FIEUWS, C. DE BLOCK, R. HILBRANDS, L. VAN HUFFEL, T. MAES, G. VANHAVERBEKE, E. DIRINCK, N. MYNGHEER, ET AL., *Comparing real-time and intermittently scanned continuous glucose monitoring in adults with type 1 diabetes (alertt1): a 6-month, prospective, multicentre, randomised controlled trial*, The Lancet, 397 (2021), pp. 2275–2283.
- [15] X. WU, S. WANG, X. WANG, AND G. HE, *A wireless real-time monitoring system on intelligent underwater pollution cleaning robot*, Journal of Computational Methods in Sciences and Engineering, 21 (2021), pp. 1051–1060.
- [16] Z. YAN, K. LI, D. SHAO, Q. SHEN, Y. DING, S. HUANG, Y. XIE, AND X. ZHENG, *Visible-light-responsive reduced graphene oxide/gc 3 n 4/tio 2 composite nanocoating for photoelectric stimulation of neuronal and osteoblastic differentiation*, RSC advances, 12 (2022), pp. 8878–8888.
- [17] G. YANG, J. WU, Q. HU, AND D. LEI, *Real-time monitoring of changes in building tilt angle using double square artificial marks*, Indoor and Built Environment, 30 (2021), pp. 777–786.
- [18] S. YANG, T. DENG, P. QIU, T. XING, J. CHENG, Z. JIN, P. LI, X. SHI, AND L. CHEN, *High-performance and stable (ag, cd)-containing znsb thermoelectric compounds*, ACS Applied Materials & Interfaces, 14 (2022), pp. 26662–26670.
- [19] L. ZHANG, D. LV, Y. GAO, AND ET AL., *Design and simulation of remote sensing monitoring method for water pollution*, Computer Simulation, 2022 (2022), pp. 289–292, 312.

*Edited by:* Sathishkumar V E

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

*Received:* May 13, 2023

*Accepted:* Jun 20, 2023







## THE OPTIMAL SELECTION OF IDEOLOGICAL AND POLITICAL ISSUES IN BUSINESS COURSES BASED ON SWARM INTELLIGENCE ALGORITHM

**Abstract.** The current optimal selection matrix for ideological and political issues of business courses is mostly set as a single objective form, and the topic selection is limited in scope, increasing the mutation rate of the optimal selection of topics. Therefore, the design and analysis of the optimal selection method for ideological and political issues of business courses based on the swarm intelligence algorithm is proposed. According to the actual measurement needs and standards, extract the optimal characteristics of ideological and political issues selection of the curriculum, use a multi-level approach to break the limits of issues selection, establish a multi-level swarm intelligence selection matrix, build an optimal selection model for ideological and political issues of business and trade courses based on swarm intelligence accounting, and achieve the optimal selection of issues through group fixed-point optimization. The novelty of this work lies in the design and analysis of the optimal selection method for ideological and political issues in business courses using a swarm intelligence algorithm. This approach introduces a new way of selecting topics by harnessing the power of collective intelligence inspired by the behavior of insects or animals. The final test results show that the mutation rate of the optimal selection of the three topics finally screened using the swarm intelligence algorithm is better controlled below 0.2 through the measurement of five classes, indicating that the topic is more practical, more targeted, and has better discussion value.

**Key words:** Swarm intelligence algorithm; Commerce and trade; Professional courses; Curriculum ideological and political; Ideological and political issues; Optimal solution;

**1. Introduction.** The selection of ideological and political issues for business courses is relatively complex and complicated, and many factors need to be considered in the design process, such as market price changes, product quality, quantity, and business trends [21]. Therefore, only by integrating the professional knowledge and practice of business courses with ideological and political issues, can we further understand the discipline and teaching characteristics of the issues, use special carriers and channels to improve and optimize the issues, adjust and modify the problems of the process of ideological and political teaching, and gradually refine the content of ideological and political teaching. Give play to the most significant teaching function and learning value of the topic [18]. The traditional optimal selection method of ideological and political issues in business courses is mostly a targeted form of processing. Although it can achieve the expected teaching tasks, due to the variability and transformation of business courses, the teaching objectives of ideological and political issues in the courses are often unstable, resulting in students being restricted by various factors and environments when learning. Finally, the optimal solution [5] could not be determined. Therefore, this paper proposes the design, verification, and analysis of the optimal selection method of ideological and political issues for business professional courses based on swarm intelligence algorithm. The so-called swarm intelligence algorithm mainly refers to a dynamic computing method [6] that uses the simulation of insects or animals, takes the population as the main environment for measurement, filters and eliminates problems one by one in the form of cooperation, and finally obtains the optimal solution. Integrating this algorithm with the optimal selection of ideological and political issues of business professional courses can further expand the optimal selection range of actual ideological and political course issues to a certain extent, and accurately obtain the optimal value [22] of each issue selection stage through the algorithm. The setting of issues also needs to take students as the main body, follow the personalized requirements, and gradually build a more diversified and flexible issue selection structure in combination with multi module and multi-objective business professional knowledge, so as to ensure that the ideological and political curriculum itself is comprehensive and integrated, and solve the existing teaching problems faster and better in the complex issue environment, form multi-level ideological and political teaching objectives [7]. Take business professional courses as the basis of teaching, and slowly implement the students' own core literacy. At the same time, with the help and support of swarm intelligence algorithm, it is also necessary to further optimize and improve the initial optimal selection mode of ideological and political issues

of the curriculum [23]. We screened and approved professional topics from various aspects, and integrated business professional knowledge related to business courses, catering to the development trend of multi-source education, and providing reference and reference for innovation and research in the selection of ideological and political topics in subsequent courses [1].

Political issues that can be incorporated into business courses:

1. Analyzing the impact of government policies and regulations on business operations, such as taxation policies, trade agreements, environmental regulations, and labor laws. This includes understanding how businesses navigate the political landscape and adapt their strategies accordingly.
2. Examining businesses' ethical and social responsibilities towards stakeholders and society at large. This includes exploring sustainability, corporate governance, diversity and inclusion, and philanthropy.
3. Investigating the political implications of globalization on business activities, including international trade, foreign investment, and cultural integration. This involves understanding the role of international organizations, regional blocs, and geopolitical factors in shaping global business environments.
4. Exploring ethical dilemmas and issues related to corruption in business practices. This includes studying cases of corporate misconduct, bribery, conflicts of interest, and ethical decision-making frameworks.
5. Analyzing the role of economic policies in influencing business cycles, economic growth, and market dynamics. This includes studying fiscal and monetary policies, inflation, unemployment, and the role of central banks.
6. Investigating the influence of corporations on the political process, including lobbying, campaign financing, and corporate involvement in policymaking. This involves examining the potential conflicts of interest and their implications for democratic governance.
7. Exploring issues related to social justice, inequality, and business practices. This includes examining topics such as fair trade, income distribution, resource access, and social entrepreneurship.
8. Examining the political dimensions of emerging technologies, such as data privacy, cybersecurity, artificial intelligence, and surveillance. This involves understanding the role of governments in regulating technology and balancing innovation with social concerns.

The research work addresses optimal selection by contributing,

1. The use of a swarm intelligence algorithm in selecting ideological and political issues for business courses is a novel approach. By emulating a population's collective behavior and cooperation, the algorithm enables the identification of optimal solutions that may not be achievable through traditional single-objective selection methods.
2. The work proposes a multi-level approach to break the limitations of topic selection in business courses. By establishing a multi-level swarm intelligence selection matrix, the scope of topic selection is expanded, allowing for a more comprehensive and diverse range of issues to be considered.
3. The development of an optimal selection model tailored explicitly for ideological and political issues in business courses enhances the precision and effectiveness of the selection process. By integrating swarm intelligence accounting and employing group fixed-point optimization, the model ensures that the selected topics have practicality, relevance, and valuable discussion potential.
4. The work addresses the issue of high mutation rates in topic selection by utilizing the swarm intelligence algorithm. Through extensive testing and measurement, the study demonstrates that the algorithm effectively controls the mutation rate of the selected topics, resulting in more targeted and valuable topics for discussion.

## **2. Optimal selection method of group intelligent accounting for ideological and political issues in designing professional courses of commerce and trade.**

### **2.1. Extract the optimal characteristics of curriculum ideological and political issues selection.**

The combination of business professional teaching and curriculum ideological and political education can be designed from a wide range of perspectives, and usually will not be subject to greater restrictions [2]. However, to ensure the accuracy and reliability of the final topic selection, we can first integrate the swarm intelligence algorithm according to the actual topic direction and the ideological and political teaching content of the curriculum to extract the optimal feature of the ideological and political topic selection of the curriculum [15]. The ideological and political issues of the curriculum established this time need not only comprehensive

ideological and political knowledge [19]. We should also carry out targeted integration of business expertise [11]. Therefore, first calculate the proportion of business courses and ideological and political courses in the topic, as shown in Formula 2.1:

$$H = \frac{\Re \times \frac{1}{\kappa}}{W + \kappa(\Re - 0.2\alpha)^2} - \omega\alpha \quad (2.1)$$

In Formula 2.1,  $H$  indicate the proportion of topics,  $W$  indicates the preset coverage selection range,  $\kappa$  indicates the deviation in the selection of targeted topics,  $\Re$  represents the convertible ratio of the topic,  $\alpha$  indicates the repeating range,  $\omega$  indicates the standard value of topic selection. According to the above measurement, calculate the proportion of commercial and trade courses and ideological and political courses in the topics. When extracting the optimal selection criteria for topics, make a biased selection of materials and issues to ensure the integrity and particularity of the ideological and political topics of the curriculum [20]. Combined with the above measurement, the optimal characteristics of selecting ideological and political issues in the curriculum are extracted [24].

The first is the leading role of the curriculum's ideological and political issues [12]. In the process of learning and discussion between teachers and students, the objectives created by this topic must be highly instructive, which can include the basic teaching content, teaching activities, and teaching behavior of business and ideological and political courses, and is the most basic topic objective [4]. The second is openness. This topic must be able to ensure that students can talk freely in the learning process. The topic itself has high discussion value, wide-coverage, and strong inclusiveness, and can deepen the teaching level of the topic and sublimate the teaching content. The third is the need for mobility [13]. This part refers to the need to properly integrate some practical activities into the agenda design process, expand the optimal selection range of the agenda, increase their own research highlights, combine activities with the agenda, and improve teaching quality and effectiveness [17]. The fourth is activity, which mainly requires students and teachers to carry out effective research and exchange on the existing problems and the setting of the ideological and political direction trend of commerce, trade and curriculum in the process of topic design, so as to further deepen the topic and highlight the value of choice [14].

**2.2. Establishing a Multi level Swarm Intelligence Selection Matrix.** After completing the extraction of the optimal features for the selection of curriculum ideological and political issues, the next step is to integrate the swarm intelligence algorithm and establish a multi-level swarm intelligence selection matrix. First, set a basic discussion goal based on the actual topic selection direction, integrate the professional knowledge of business courses with the ideological and political content of the course, and first draw up the overall coverage of the topic, forming a stable and complete discussion area [16]. Then, based on this, teachers need to create a discussion environment in the classroom for the related questions. Students set the corresponding discussion level and small stage topic design goals according to the difficulty of the questions, and finally give a basic answer [9]. The whole process is actually equivalent to a matrix framework of continuous outcome problems. According to the small goals formed by the issues, the corresponding hierarchy of the matrix [3] is established. See Figure 2.1 for details.

According to Figure 2.1, complete the design and adjustment of the structure of the multi-level swarm intelligence selection matrix. At the same time, it is also necessary to establish the corresponding optimal selection criteria for topics in the matrix. Teachers should combine the professional knowledge and practical content of business courses to create a discussion situation. The topics should be closely linked with small goals, so as to further ensure that students are in discussions. In the process of analysis, further, deepen the connotation of the topic and help students gradually understand and master the deeper content of the topic [10].

However, in this part, it should be noted that when designing the matrix, students can establish corresponding selection criteria based on their own judgment and can conduct the second stage of research and discussion [8] when selecting follow-up issues. Generally, topics will not be set with a single fixed answer. They can be integrated, divergent and dialectical from multiple directions to reflect their value. Use the matrix to expand the scope of thinking and discussion of topics and deepen the theme [5].

**2.3. Building the Optimal Selection Model of Ideological and Political Issues for Group Intelligent Accounting Business Courses.** After establishing the multi-level swarm intelligence selection matrix,

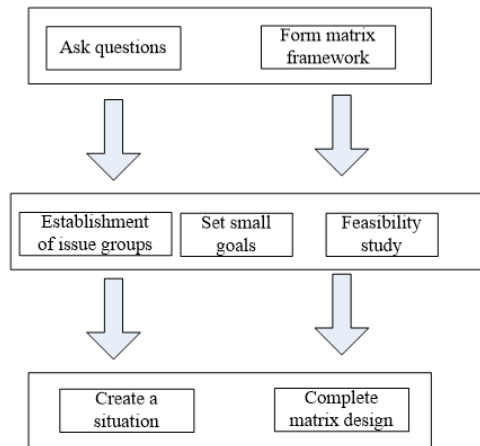


Fig. 2.1: Structure of Multi level Swarm Intelligence Selection Matrix

the next step is to integrate the algorithm to construct the optimal selection model for business courses' ideological and political issues. In fact, different from the single goal issue selection model, the two-way issue of business courses and ideological and political courses pays more attention to the choice of the issue's content and the direction of the discussion. Therefore, in establishing the model, we can first use the matrix to divide the research direction and theme of the issue, forming a structurally strong issue selection framework. Students or teachers can build a group and determine the optimal value of the issue at this time according to their own judgment and standard setting, and by integrating the swarm intelligence algorithm, as shown in Formula 2.2:

$$Y = \sqrt{\varepsilon + \sum_{v=1} \theta v - 1 \times \varepsilon \chi} \tag{2.2}$$

In Equation 2.2,  $Y$  indicates the optimal value of the issue,  $\varepsilon$  indicates the dialectical scope of the topic,  $\theta$  indicates the small goal of topic selection,  $\theta$  indicates the optional quantity,  $\chi$  indicates controllable issue deviation. According to the above settings, complete the calculation of the optimal value of the issue. In this community, the optimal value is used as the limiting selection standard to determine the mutation situation of the population. If the mutation ratio exceeds 2%, it indicates that the feasibility of this issue is not high and needs to be reset; On the contrary, if the mutation ratio does not exceed 2%, it indicates that the feasibility of this issue is high. You can set the model for comparison and selection to obtain the final optimal selection result.

**2.4. Group fixed-point optimization to achieve optimal choice of issues.** Use the above constructed optimal selection model of ideological and political issues for business and trade courses of group intelligent accounting to make basic selection of ideological and political issues for courses. Next, use group fixed-point optimization to achieve optimal selection of issues. You can first establish the corresponding screening process in the model matrix, set group selection points a, b, c, d. n, each group point is equivalent to an issue, and design specific selection principles, as shown in Figure 2.2.

According to Figure 2.2, complete the design and adjustment of the group fixed-point optimization principle. Next, according to the specific issue standards and contents of the group fixed-point, follow the initially set standards, carry out comparison and optimization, and achieve the optimal selection of issues. Matrix filtering is a digital image processing technique used to enhance or modify an image by applying a filter to its pixel values. It involves convolving a filter matrix or kernel with the image matrix, which performs a local operation on each pixel and its neighboring pixels. The filter matrix contains numerical coefficients that determine the nature of the filtering operation.

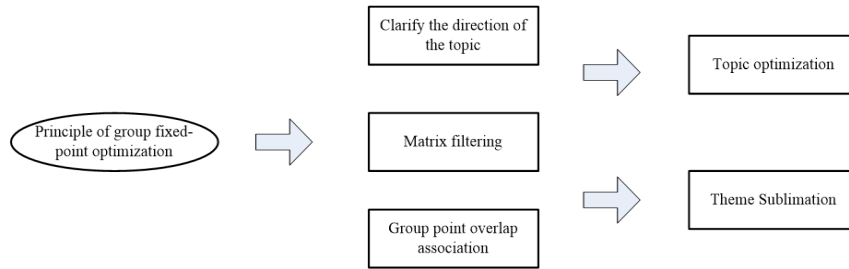


Fig. 2.2: Schematic diagram of group fixed-point optimization principle

Matrix filtering works by performing a weighted average of the pixel values within a defined neighborhood around each pixel in the image. This neighborhood is determined by the size and shape of the filter matrix. The filter matrix is usually a square matrix with odd dimensions, such as  $3 \times 3$ ,  $5 \times 5$ , or  $7 \times 7$ .

During the filtering process, the filter matrix is overlaid on the image matrix, and the corresponding elements of the two matrices are multiplied together. The resulting products are then summed to obtain a new value for the central pixel. This process is repeated for each pixel in the image, resulting in a filtered output image.

The coefficients in the filter matrix determine the influence of each neighboring pixel on the central pixel. They can be designed to achieve different filtering effects, such as blurring, sharpening, edge detection, noise reduction, and image enhancement. For example, a Gaussian filter uses a weighted average to blur the image, while a Laplacian filter enhances edges by highlighting intensity variations.

**3. Experiment.** This time is mainly to analyze and verify the practical application effect of the optimal selection method for ideological and political issues of business professional courses based on the swarm intelligence algorithm. Considering the authenticity and reliability of the final test results, the analysis is carried out by comparison, and the ideological and political issues of business professional courses at R University are selected as the actual measurement target. According to the actual selection needs and changes in the standards of the topics, the final measured results are compared and studied. Next, the relevant test environment is built by integrating the swarm intelligence algorithm.

**3.1. Experiment preparation.** Integrate swarm intelligence algorithm to build a measurement environment for the optimal selection of ideological and political issues in business professional courses of R university. Five classes are selected as the test objects in R University. The three classes are all learning classes for business majors, and the number of students in each class is about 65. Situational teaching method, discussion method, practice method, and question and answer method are mostly used in the daily teaching of ideological and political courses, and most of the courses are lectured, students' participation is not high, and their understanding and learning of knowledge points of ideological and political courses are not deep, leading to the learning effect cannot reach the expected standard. Under such a background, the best choice of ideological and political issues of the curriculum is needed. First of all, according to the actual teaching situation, the optimal selection of topics is divided into four stages, which are the determination of primary topic direction, feasibility analysis and topic selection, practical demonstration of topics, and optimal selection of topics. The specific topic definition and selection structure design are shown in Table 3.1.

According to Table 3.1, complete the structural design and research of topic definition and selection. The learning level of the selected five classes of business courses is basically the same, the student's ability is almost the same standard, and there is basically no big gap between the average scores of professional disciplines. The cycle of the topic selection test is set as one week, and every day is a test summary stage. The processing data and information of topic selection need to be recorded for future use. On this basis, 10 initial curriculum ideological and political issues are selected as the objectives of the measurement. At this time, combined with the needs of teaching, a practical issue selection standard is formulated to gradually form a complete

Table 3.1: Structure setting table of topic definition selection

Topic selection stage	Topic definition	Execution form	Replenish
Determination of basic topic direction	According to the professional knowledge and practice of business courses, determine the specific research direction of curriculum ideological and political	Group demonstration	Multiple topics can be determined for comparison and selection
Determination of basic topic direction	Determine whether the selected topics have real meeting and sub-research value, and conduct equivalent screening	Group demonstration	Unify feasibility and screening criteria
Practical demonstration of issues	Conduct practical demonstration and analysis on the selected topics	Personal argument	Build practical solutions
Optimal choice of issues	Select the best topic	Group demonstration	Calculate the optimal solution

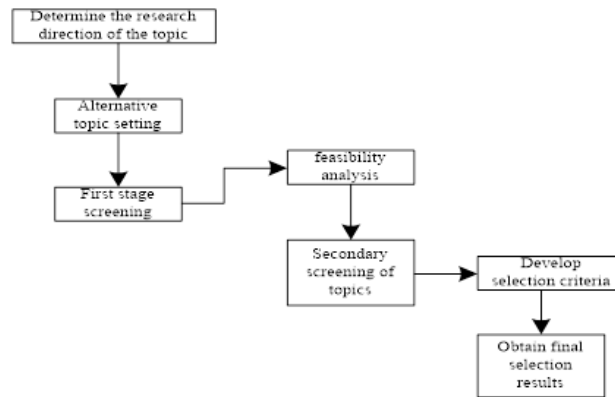


Fig. 3.1: Structure diagram of topic selection of ideological and political courses

optimal selection structure of curriculum ideological and political issues, so as to achieve the establishment of a basic testing environment. Next, specific tests and analyses are conducted by integrating swarm intelligence algorithms.

**3.2. Experimental process and result analysis.** In the above-built test environment, the swarm intelligence algorithm is integrated, and then specific verification and analysis are carried out. Let the three selected classes discuss together to determine the research direction of one or more business courses' ideological and political issues, and choose the one that is most interesting to students and has the highest teaching value. After clarifying the direction of the topic, students from five classes need to build multiple alternative topic plans for that direction. It should be noted that each topic is independent and has its own research and discussion value. On this basis, the corresponding structure is established based on the selection criteria of ideological and political issues of the curriculum, as shown in Figure 3.1.

According to Figure 3.1, complete the design and research on the topic selection of ideological and political

Table 3.2: Formulation of optimal selection criteria for ideological and political issues of curriculum

Division of topic selection direction	Proportion of topics	Evaluation and selection criteria
Business courses	Discipline literacy accounts for 0.6	Issue setting
		Topic selection
Curriculum ideological and political	Ideological and political core literacy accounts for 0.4	Scenario setting and arrangement
		Optimal selection activity design

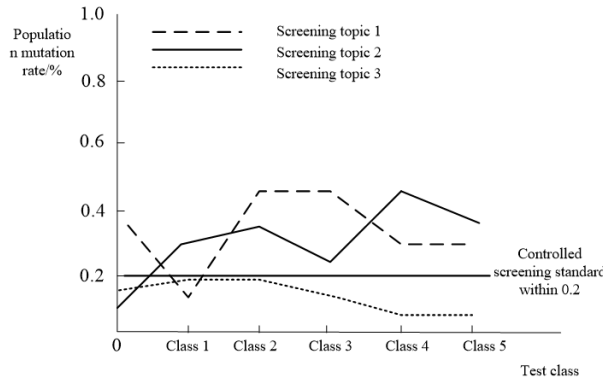


Fig. 3.2: Comparison and Analysis of Test Results

courses. At this time, in this stage, the ideological and political issues of business courses have been screened to some extent. Next, we need to combine the swarm intelligence algorithm to select the optimal issues and formulate corresponding dynamic selection criteria, as shown in Table 3.2.

According to Table 3.2, complete the formulation and adjustment of the optimal selection criteria for curriculum ideological and political issues. On this basis, the topics finally selected are analyzed. At this time, the group intelligence algorithm is used to set the coverage of the topics as a controllable group. Once the ideological and political issues of the curriculum are different from the group of business courses, the mutation rate will increase. Select the three topics finally selected for measurement and comparison, synthesize the data obtained and the information of related topics, and comprehensively evaluate to calculate the mutation rate of topic selection, as shown in Formula 3.1:

$$G = \sum_{y=1}^{\mathfrak{S}} \mathfrak{S}y \times \alpha^2 + (1 - \mathfrak{S}) \tag{3.1}$$

In Formula 3.1,  $G$  indicates the mutation rate of the population selected for the topic,  $\mathfrak{S}$  indicates the scope of discussion,  $y$  indicates the number of optional topics,  $\alpha$  indicates the population transition area. According to the above determination, the analysis of the test results can finally be completed, as shown in Figure 3.2.

The results of the analysis and verification of the test demonstrate the effectiveness of the swarm intelligence algorithm in achieving an optimal selection of ideological and political issues for business courses. Figure 3.2 illustrates the outcomes of the test and provides valuable insights. The measurement of five classes was conducted to evaluate the mutation rate of the optimal selection for the three topics that were finally chosen. The mutation rate refers to the degree of change or deviation from the original selection. In this context, a lower mutation rate indicates that the selected topics have a higher level of practicality, relevance, and discussion value.

Based on the results shown in Figure 3.2, it can be observed that the swarm intelligence algorithm successfully controlled the mutation rate of the optimal topic selection below 0.2. This indicates that the selected topics remained stable and consistent throughout the measurement process, with minimal deviation or variation from the initial selection. The controlled mutation rate suggests that the selected topics are more practical and targeted. The algorithm was able to identify topics that align closely with the learning objectives and requirements of the business courses. These topics are expected to provide meaningful and valuable discussions for students, enhancing their understanding and engagement with the course material.

Furthermore, the results indicate that the selected topics have better discussion value. This implies that the topics are not only relevant and practical but also stimulate productive and insightful discussions among students. The inclusion of these topics in the curriculum can foster critical thinking, promote interdisciplinary perspectives, and encourage active participation and exchange of ideas among students.

**4. Conclusion.** To sum up, the above is the design, verification and analysis of the optimal selection method for ideological and political issues of business professional courses based on swarm intelligence algorithm. Compared with the initial issue selection form, this time, with the help of swarm intelligence algorithm, the actual issue selection form constructed is relatively more stable, diversified, and targeted, which can ensure the accuracy and reliability of the issue results. With this colleague, the process of the topic is also more complete, which makes it easy to form an optimal knowledge framework, gradually combining the teaching background materials and content, and relying on situational materials in the ideological and political curriculum, further expanding the discussion and teaching of the ideological and political topic of the curriculum, simplifying the selection process within a reasonable range, strengthening the depth of the topic, and giving better play to the teaching value.

However, it is important to acknowledge certain limitations of this study. Firstly, the research focused on the application of the swarm intelligence algorithm in the selection of ideological and political issues in business courses, without considering the potential challenges or drawbacks of this approach. Additionally, the study did not explicitly address the perspectives and preferences of students, which could have an impact on the effectiveness of the selected topics. Future research should consider these limitations and further investigate the potential limitations and benefits of the swarm intelligence algorithm in selecting ideological and political issues. Moreover, incorporating student feedback and engagement in the selection process can provide valuable insights for improving the overall effectiveness of the curriculum.

#### REFERENCES

- [1] H. AI, *The influence of new media on college students' ideological and political education and the countermeasures*, Journal of Contemporary Educational Research, 5 (2021), pp. 44–46.
- [2] R. G. D. ALMEIDA, *A macroeconomic view of public choice: new political macroeconomics as a separate tradition of public choice*, 2021.
- [3] A. AWAD, R. SALEM, H. ABDELKADER, AND M. ABDSALAM, *A swarm intelligence-based approach for dynamic data replication in a cloud environment*, Int. J. Intell. Eng. Syst, 14 (2021), pp. 271–284.
- [4] J. CHANG, Y. LI, AND H. ZHENG, *Research on key algorithms of the lung cad system based on cascade feature and hybrid swarm intelligence optimization for mkl-svm*, Computational Intelligence and Neuroscience, 2021 (2021).
- [5] T. CHEN, *Research on the dilemma and breakthrough path of ideological and political education in colleges and universities in the era of big data*, Journal of Higher Education Research, 3 (2022), pp. 203–206.
- [6] P. CHENG, L. YANG, T. NIU, AND ET AL., *On the ideological and political education of material specialty courses under the background of the internet*, Journal of Higher Education Research, 3 (2022), pp. 79–82.
- [7] Y.-N. DIAO, Y. HAN, AND X. ZENG, *Investigation and analysis of children's choice of environment in oral clinics*, Hua xi kou Qiang yi xue za zhi= Huaxi Kouqiang Yixue Zazhi= West China Journal of Stomatology, 39 (2021), pp. 182–187.
- [8] S. DWIVEDI, M. VARDHAN, S. TRIPATHI, AND A. K. SHUKLA, *Withdrawn: Scaling up detection rate in intrusion detection using self-adaptive swarm intelligence mechanism*, 2021.
- [9] A. A. EL-SALEH, T. M. SHAMI, R. NORDIN, M. Y. ALIAS, AND I. SHAYEA, *Multi-objective optimization of joint power and admission control in cognitive radio networks using enhanced swarm intelligence*, Electronics, 10 (2021), p. 189.
- [10] E. EL-SHAFEIY, K. M. SALLAM, R. K. CHAKRABORTY, AND A. A. ABOHANY, *A clustering based swarm intelligence optimization technique for the internet of medical things*, Expert Systems with Applications, 173 (2021), p. 114648.
- [11] C. JIN, J. LI, Y. XI, T. HAN, S. LIU, Y. LUO, R. CAO, AND G. SHEN, *Problems and solutions in ideological and political education under the training mode of excellent and innovative talents in colleges and universities*, Asian Agricultural Research, 14 (2022), pp. 51–57.



- [12] Q. LI AND S. LI, *Optimization of artificial cnn based on swarm intelligence algorithm*, Journal of Intelligent & Fuzzy Systems, 40 (2021), pp. 6163–6173.
- [13] M. MANSOOR, A. F. MIRZA, F. LONG, AND Q. LING, *An intelligent tunicate swarm algorithm based mppt control strategy for multiple configurations of pv systems under partial shading conditions*, Advanced Theory and Simulations, 4 (2021), p. 2100246.
- [14] W. K. MASHWANI, R. HAIDER, AND S. BRAHIM BELHAOUARI, *A multiswarm intelligence algorithm for expensive bound constrained optimization problems*, Complexity, 2021 (2021), pp. 1–18.
- [15] Y. NIKIFOROV, *The ideological and political support for the transformation of the economy and society in the ussr of the 1950–1980s: historical challenges and hidden trends*, Vestnik of Kostroma State University, 26 (2021), pp. 50–57.
- [16] C. NTAKOLIA AND D. K. IAKOVIDIS, *A swarm intelligence graph-based pathfinding algorithm (sigpa) for multi-objective route planning*, Computers & Operations Research, 133 (2021), p. 105358.
- [17] A. P. PANDIAN, *Novel distance estimation based localization scheme for wireless sensor networks using modified swarm intelligence algorithm*, IRO Journal on Sustainable Wireless Systems, 2 (2021), pp. 171–176.
- [18] X. TAN, *An interpretation of the ideological and political construction of college courses from the perspective of reception aesthetics\**, Psychology Research, 12 (2022), pp. 840–844.
- [19] H. TANG AND Q. SHI, *On the ways to improve the affinity of ideological and political education psychology from the perspective of teachers*, Psychology of China, 3 (2021), pp. 650–655.
- [20] X.-L. WANG, J.-W. LU, G. XIAN-GUANG, AND S.-Q. YANG, *A global synthesis approach for optimizing the meshing performance of hypoid gears based on a swarm intelligence algorithm*, Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science, 235 (2021), pp. 1368–1388.
- [21] X. WU, *Research on the reform of ideological and political teaching evaluation method of college english course based on “online and offline” teaching*, Journal of Higher Education Research, 3 (2022), pp. 87–90.
- [22] A. XU, *A new solution concerning choice-of-law for the assignment of debts*, International & Comparative Law Quarterly, 70 (2021), pp. 665–696.
- [23] Y. ZHANG, *Research on ideological and political education of college students from the perspective of system theory*, Frontier of Higher Education, 2 (2021), pp. 1–17.
- [24] W.-S. ZHAO, B.-X. WANG, D.-W. WANG, B. YOU, Q. LIU, AND G. WANG, *Swarm intelligence algorithm-based optimal design of microwave microfluidic sensors*, IEEE Transactions on Industrial Electronics, 69 (2021), pp. 2077–2087.

*Edited by:* Sathishkumar V E

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

*Received:* May 14, 2023

*Accepted:* Jul 9, 2023





## INTELLIGENT ELEMENT LAYOUT ALGORITHM OF CERAMIC DESIGN DRAWING BASED ON GENETIC ALGORITHM

**Abstract.** Conventional intelligent layout algorithm of ceramic design drawing elements mainly uses density distribution method to obtain ceramic design drawing element units, which is vulnerable to the influence of pattern density, resulting in a low distribution balance of design drawing elements. Therefore, a new intelligent layout algorithm for ceramic design drawing elements needs to be designed based on a genetic algorithm. That is, combining the characteristics of ceramic design drawing elements, the intelligent parting structure of ceramic design drawing elements is constructed, and then the brilliant layout model of ceramic design drawing elements is designed using genetic algorithm. The experimental results show that the developed intelligent layout algorithm of ceramic design elements has a high distribution balance, which proves that the designed intelligent layout algorithm of ceramic design elements has good performance, reliability, and certain application value, and has made specific contributions to improving the aesthetic quality of ceramic design drawings.

**Key words:** Genetic algorithm; Ceramics; Design chart; Element; Intelligence; Layout diagram

**1. Introduction.** Ceramics is a typical handicraft mainly composed of pottery and porcelain. Ceramic has a long history in China and is also one of the essential representatives of Chinese art [7]. As early as the Neolithic Age, China began to study the style of ceramic display and try to combine the pattern elements contained in it. By the Ming and Qing Dynasties [12], China's ceramic production technology was perfect and various patterns were lifelike. In recent years, the living standard of our country has gradually improved, and the selection of ceramics is not only focused on its use [9], but also the graphic element composition (aesthetic visual characteristics) of its appearance has become the main consideration for people to select ceramics.

The research shows that ceramics show strong irregularity [26] due to its special structure, so when designing patterns, they pay special attention to the intelligent layout and balance elements of graphic elements. To improve the layout balance of graphic elements, intelligent design [18] must be carried out in combination with the original structural characteristics of ceramics, to combine the practicality and appearance of ceramics organically. The early layout method of graphic elements originated from natural factors. People often carry out artistic visual design [10] on ceramics according to the unique perception of nature to improve the comprehensive aesthetic sense of ceramic patterns. In addition, some ceramics pay attention to the combination of information display and transmission in the process of graphic element design to further display the aesthetic characteristics of ceramics.

Ceramics have essential application value in China [16], and are also used in exhibitions, daily use, and many other aspects. In recent years, the competition in the ceramic market has gradually intensified. Suppose you want to improve your comprehensive strength. In that case, you must ensure the balance of ceramic patterns, make them meet people's aesthetic needs, and ensure that ceramic patterns have effective aesthetic significance. In the computer age, ceramic graphic element design technology is undergoing a significant transformation, gradually changing from early manual design to intelligent design using computers [1]. Although intelligent design can improve the efficiency of pattern design and reduce the design cycle, it often fails to consider the actual needs of ceramic graphic element design, resulting in a single design pattern and rigid layout structure. Relevant researchers have designed several conventional ceramic visual element intelligent layout algorithms [15] according to the smart layout characteristics of graphic elements, but the balance of these algorithms is relatively poor. In order to solve the above problems and improve the aesthetic balance of ceramic design drawing elements, this paper designs a new intelligent layout algorithm of drawing elements based on a genetic algorithm.

The conventional intelligent layout algorithms for ceramic design drawing elements primarily rely on density distribution methods to obtain ceramic design element units. However, these methods are susceptible to the

influence of pattern density, resulting in a low distribution balance of design drawing elements. Therefore, there is a need to develop a new intelligent layout algorithm that addresses these limitations and improves the overall distribution balance of ceramic design drawing elements. Genetic algorithms offer a promising approach to optimize the arrangement of these elements, considering their specific characteristics and achieving better aesthetic quality in ceramic design drawings.

Research Question are as follows:

1. How can the distribution balance of ceramic design drawing elements be improved by designing an intelligent layout algorithm based on a genetic algorithm?
2. How can the intelligent parting structure of ceramic design drawing elements be constructed by incorporating their specific characteristics?
3. How does the proposed algorithm contribute to improving the aesthetic quality of ceramic design drawings?
4. What are the potential applications and practical value of the designed intelligent layout algorithm in the field of ceramic design.

## **2. Design of intelligent layout algorithm of ceramic design drawing elements based on genetic algorithm.**

**2.1. Build intelligent parting structure of ceramic design drawing elements.** To solve the problem that the element distribution balance of the ceramic design drawing is reduced due to the influence of the pattern density when the density distribution method is used to obtain the element units of the ceramic design drawing, this paper constructs the intelligent parting structure of the ceramic design drawing elements based on the characteristics of the ceramic design drawing elements. In the conventional layout of graphic elements, several elements [6] are constantly added to form a variety of layout styles, or the same elements can be arranged in equal proportion to create a regular layout. On this basis, to effectively build the intelligent parting structure of ceramic design drawing elements, it is necessary to design the corresponding intelligent layout plug-in [7], adjust the brilliant layout rule library, save it in the template, and intelligently generate drawing elements. In the above process, you can save the sketch simulation using the same collection class to obtain intelligent parting parameters. Based on the above steps, you need to load the initial diagram elements first, and the loading process is shown in Figure 2.1.

It can be seen from Figure 2.1 that an intelligent typing record will be generated when loading the initial diagram elements. At this time, it can be recorded in the rule base, to record the initial position [8, 24] displayed by the diagram elements for a balanced layout of the diagram elements.

After the initial element is loaded, it is necessary to determine the basic loading information of the element, further obtain the element address, and ensure that the element has not been removed, and there is relevant graphic data [22] in the graphic element library. The corresponding management rules [4] are followed during the basic information processing of diagram elements, so a basic sketch template can be generated to realize an intelligent display of parameters. At this time, the schematic diagram of diagram element information saving rules is shown in Figure 2.2.

It can be seen from Figure 2.2 that the intelligent layout threshold can be set in combination with the above graph element information saving rules, but if you want to improve the intelligent layout effect, you need to save the graph element intelligent generation parameter [23] to obtain the original layout template. The intelligent parting structure of ceramic design drawing elements has several basic characteristics. First, it has self-similarity, that is, there is a similarity relationship between the design part and the design subject. Second, it is self-affine, that is, it carries out similarity expansion [5]. Finally, it needs to determine the parting dimension to generate an effective intelligent parting structure. At this time, the generated parting structure is shown in Figure 2.3.

It can be seen from Figure 2.3 that using the above intelligent layout and typing structure of ceramic design drawing elements can determine the basic design characteristics of the drawing elements, and carry out the metrological design into processing [19], so as to generate a new subject fractal diagram element design graph, and ensure the balance of intelligent layout to the greatest extent.

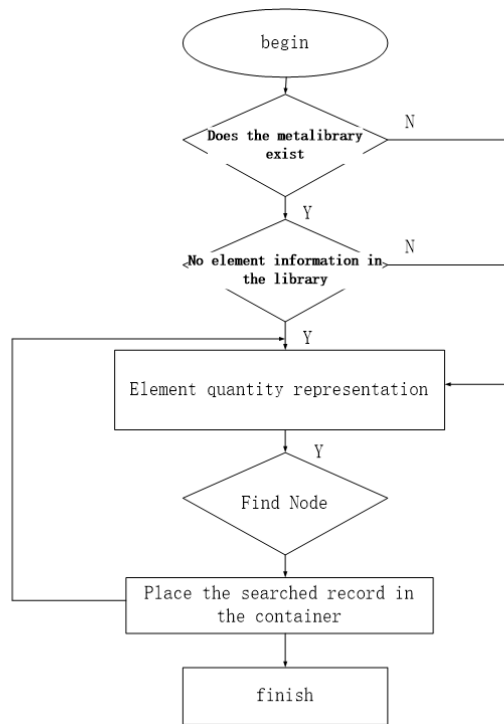


Fig. 2.1: Initial diagram element loading process

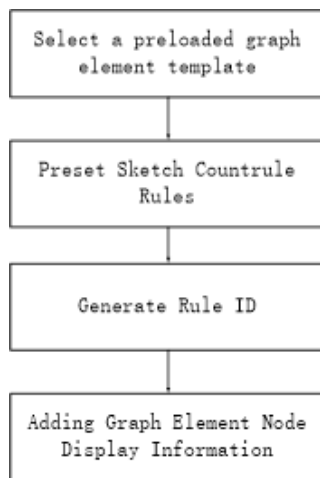


Fig. 2.2: Rules for saving diagram element information

**2.2. Design of intelligent layout model of ceramic design drawing elements based on genetic algorithm.** Genetic algorithms can combine biological evolution law to conduct random searches and comprehensively search for the specified graph elements, thus improving the effectiveness of the intelligent layout of graph elements. First, ceramic design diagram elements can be used as the genetic basis to generate the initial genetic mechanism and build an effective iterative process [25]. In this process, there are fitness differences for

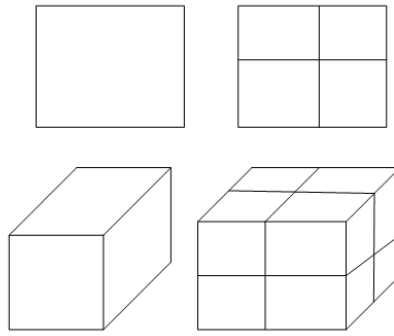


Fig. 2.3: Intelligent layout and parting structure of ceramic design elements

different populations, and the heredity at this time is random. The initial population generated at this time  $p(0)$  as shown in 2.1.

$$p(0) = \{x_1^0, x_2^0, \dots, x_N^0\} \tag{2.1}$$

In formula 2.1,  $x_1^0, x_2^0, \dots, x_N^0$  represent the initial population individuals respectively, and the keyword vector at this time is shown in 2.2.

$$c = \langle c_1, c_2, \dots, c_n \rangle \tag{2.2}$$

In formula 2.2,  $c_1, c_2, \dots, c_n$  represents an internal keyword [17]. Different intelligent layout weights have floating point differences, and the main data structures of genetic algorithms are also different. Therefore, it is necessary to construct a descending vector according to the average weight of keywords, and perform genetic operations on the initial population, that is, select the primitive individuals [11] that need to be copied, and ensure that they can enter the next generation of genetics based on fitness, so as to carry out effective design [21, 14]. Genetic approximation at this time  $m(H, t + 1)$  as shown in 2.3.

$$m(H, t + 1) = \frac{f(H)}{f} \left[ 1 - \frac{P \cdot \delta(H)}{L - 1} \right] OH \tag{2.3}$$

In formula 2.3,  $f(H)$  represents the genetic selection operator definition moment [13],  $f$  represents low order genetic parameters,  $P$  represents the genetic index,  $\delta(H)$  represents the variation selection value,  $L$  represents near genetic weight,  $OH$  represents the layout equalization parameter [3], and the binary coding accuracy at this time  $\xi$  as shown in 2.4.

$$\xi = \frac{U - U_A}{L - 1} \tag{2.4}$$

In formula 2.4,  $U$  represents the coding constant [2],  $U_A$  representing the individual code length, the above intelligent layout based on genetic algorithm can be used to build an effective intelligent layout model of ceramic design drawing elements  $Q$ , as shown in 2.5.

$$Q = \frac{1}{m} sim(w, d) \tag{2.5}$$

In formula 2.5,  $m$  represents the number of individuals arranged,  $sim(w, d)$  graph element keyword vector [20], using the above ceramic design graph element intelligent layout model can effectively carry out comprehensive optimization to ensure the reliability of graph element layout.

**3. Experiment.** In order to verify the comprehensive performance of the designed intelligent layout algorithm of ceramic design drawing elements based on genetic algorithm, this paper built an experimental platform and compared it with the conventional intelligent layout algorithm of ceramic design drawing elements. The experiments are as follows.

**3.1. Preparation.** In combination with the demand for the intelligent layout of ceramic design elements, this paper selects P2P as the experimental platform for the intelligent layout of ceramic design elements. This platform mainly uses IM for collaborative development, which can generate basic ceramic printing patterns, ceramic moire patterns, and ceramic basic deformation patterns. In order to improve the effectiveness of the experiment, the platform uses a Web server as the experimental client. In addition, the experimental platform can display design resources in real-time and realize the sharing of basic primitive resources.

In this research, the paper selects a P2P (Peer-to-Peer) experimental platform for the intelligent layout of ceramic design elements. The choice of the P2P platform is motivated by the need for collaborative development and efficient generation of various ceramic patterns and deformations. The platform utilizes instant messaging (IM) for collaborative development, enabling the generation of basic ceramic printing patterns, ceramic moire patterns, and ceramic basic deformation patterns.

The utilization of IM in the experimental platform allows for real-time collaboration and sharing of design resources among multiple users. This feature enhances the effectiveness of the experiment by enabling efficient communication and coordination among designers and researchers involved in the ceramic design process. By utilizing a P2P framework, the platform enables decentralized and distributed sharing of design resources, which facilitates seamless collaboration and reduces the potential bottlenecks associated with traditional client-server architectures.

To enhance the experimental process, a web server is employed as the experimental client within the P2P platform. This choice allows for a user-friendly interface and facilitates the real-time display of design resources. The web-based client provides a convenient and accessible environment for designers to interact with the intelligent layout algorithm and explore the generated ceramic design elements. It allows users to visualize the design resources and make adjustments or modifications as needed.

Moreover, the experimental platform supports real-time sharing of basic primitive resources. This feature enables designers to access and utilize a diverse range of foundational ceramic design elements, such as basic printing patterns or deformations, provided by the platform. By incorporating shared resources, designers can leverage existing designs as building blocks to create novel and unique ceramic design compositions. The collaborative aspect of the platform allows designers to contribute their own resources, fostering a collaborative and creative design environment.

The selection of the P2P experimental platform, utilizing IM for collaboration, a web server as the experimental client, and real-time sharing of design resources, addresses the requirements for an efficient and collaborative environment in the intelligent layout of ceramic design elements. These features enhance the experimental process, facilitate effective communication and resource sharing, and contribute to the overall advancement of ceramic design practices. The selected experimental platform architecture is shown in Figure 3.1.

It can be seen from Figure 3.1 that the above experimental platform meets the requirements of intelligent layout experiment of graph elements. In order to improve the extraction efficiency of experimental data, this paper selects ASP as the experimental support, which can effectively manage the experimental cases and ensure the accuracy of the experimental structure. In addition, the above experimental platform has added a safety management part, which can be used for experimental authentication. Before the experiment, it is necessary to preset the layout area. This paper selects CAD software to draw the basic practical diagram elements, some of which are shown in Figure 3.2.

It can be seen from Figure 3.2 that the above graphic elements output by CAD has certain regularity. You can select the corresponding experimental menu through the main interface to process the parting matrix and edit the details to the specified format. Currently, the parameters of each basic graphic element are shown in Table 3.1.

It can be seen from Table 3.1 that the above graph elements can be uniformly designed with the specified mathematical description curve to obtain the basic experimental pixels and adjust the intelligent layout definition of graph elements. There are certain differences in the comprehensive processing of different graph

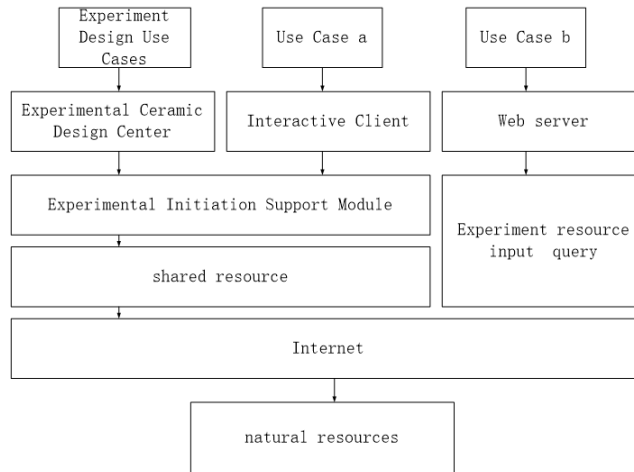


Fig. 3.1: Experimental platform architecture

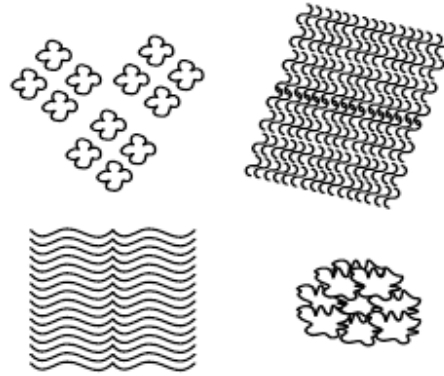


Fig. 3.2: Basic diagram elements

Table 3.1: Basic drawing element parameters

Figure Element Type	Resolution (PPI)	Image combination data amount (bit)	Basic composition parameters
Figure Element Group Type	72	3654412	1.8445
Graph element separation	72	4586635	1.9845
Figure Element Contact	72	4186632	1.2265
Figure Element Overlay	90	6254641	1.7471
Figure Element Overlay	90	6888122	1.4554
Graph element combination	90	6235141	1.9963
Figure Element Clipping	72	4584112	1.8874
Diagram Element Overlay	72	3565458	1.2252
Coincidence of graph elements	90	6845412	1.8543
Figure Element Rotation	90	6564824	1.6352
Figure element translation	90	6888452	1.8896



Table 3.2: Percentage of Element Thickness Pattern in Ceramic Design Drawing

Thickness	Thickness equalization coefficient	Figure Element Distribution Percentage (%)	Component ratio (%)	Composition mode Composition mode
0.2E-03	0.125	58.51	45.69	Figure Element Shape and Shape Combination
0.06E-03	0.833	41.36	44.32	There is no contact between shapes and there is a certain distance between them
0.06E-02	0.758	10.25	85.51	The edges between the shapes are exactly tangent to each other
0.03E-03	0.521	11.42	15.58	The relationship between shapes is overlapping
0.2E-02	0.466	5.62	85.33	The overlapping of transparency between shapes
0.6E-03	0.585	3.35	19.54	Shape combines with shape to form a new shape
0.06E-03	0.647	1.40	23.55	The areas covered by shapes and shapes are cut off
0.2E-03	0.869	2.98	36.95	Overlap each other
0.13E-03	0.853	3.85	29.54	Mutual coincidence
0.5E-03	0.652	14.98	22.71	Composite rotation
4E-02	0.754	15.22	20.56	Composite rotation
0.06E-03	0.658	24.96	36.85	Multiple combinations

elements. Bitmap elements need to be smoothed during processing, sketch elements need to use arrows for intuitive conversion, and rotating graph elements need to use Get Pixel to determine the color of graph elements before and after rotation, plan the central coordinates, and obtain the resolution of unit graph elements. For a given bitmap element, XML needs to be used to effectively describe the graph element type during the scaling operation, generate the number of basic graph element objects, and obtain the corresponding relationship between the experimental graph elements. After the basic drawing elements are processed, this paper selects the distribution balance index of design drawing elements  $r$  as an experimental index, the calculation formula is as follows 3.1.

$$r = \frac{F}{D} \quad (3.1)$$

In formula 3.1,  $F$  represent the distribution position of design drawing elements,  $D$  represents the preset distribution position of drawing elements. The higher the distribution balance of design drawing elements, the better the intelligent layout effect of drawing elements. On the contrary, the lower the distribution balance of design drawing elements, the poorer the intelligent layout effect of design drawing elements. The basic attributes of the diagram elements are different, but they are all determined by the Element tag, and do not include child nodes and corresponding attributes. Therefore, during the operation of the diagram elements, it is necessary to carry out standardized deletion according to the actual situation of Doc, determine the location of the Microsoft XML Parser experimental document, and effectively encapsulate the experiment, so as to create a basic document object and reasonably add the design root node.

Ceramic pattern design has an important relationship with the actual structure of ceramics. Therefore, the uniformity coefficient can be preset according to the thickness of different ceramic structures to determine the distribution percentage of diagram elements, as shown in Table 3.2.

It can be seen from Table 3.2 that the experimental gradient layer can be effectively planned by combining the thickness pattern percentage of the ceramic design elements in Table 3.2, so as to preset the initial distribution position of the diagram elements. After setting the balance coefficient of the intelligent layout of the above diagram elements, this experiment is based on Web2.0 and uses VC++60 as the development language to develop the Microsoft Windows 2003 standard experimental operating environment. At this time, it needs

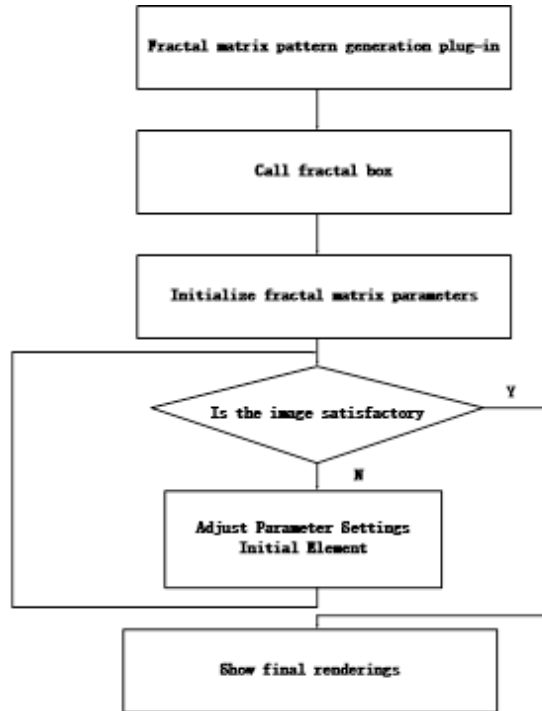


Fig. 3.3: Intelligent layout and typing process

to ensure that the server is Microsoft Sql server2000 and the browser needs to select Microsoft IE 5.0. At this time, you can set the intelligent layout classification process, as shown in Figure 3.3.

It can be seen from Figure 3.3 that combining the above generated intelligent layout and typing process can effectively process the elements of the experimental diagram and carry out subsequent intelligent layout experiments.

**3.2. Experimental results and discussion.** Combined with the above experimental preparations, the intelligent layout experiment of ceramic design diagram elements can be carried out, that is, the intelligent layout algorithm of ceramic design diagram elements based on genetic algorithm and the conventional intelligent layout algorithm of diagram elements designed in this paper can be used to arrange in the built experimental platform, and formula 3.1 can be used to calculate the distribution equilibrium index of diagram elements obtained by the two methods under different thickness structures. The experimental results are shown in Table 3.3.

It can be seen from Table 3.3 that the intelligent layout algorithm of ceramic design drawing elements designed in this paper based on genetic algorithm has a high index of element distribution balance under different thickness structures, while the traditional intelligent layout algorithm of ceramic design drawing elements has a relatively low index of element distribution balance under different thickness structures. It proves that the intelligent layout algorithm of ceramic design drawing elements based on genetic algorithm designed in this paper has good performance, reliability and certain application value.

**4. Conclusion.** Porcelain has a history of mailing in China. It is a common handicraft. In recent years, with the improvement of people's aesthetic appreciation, the choice of ceramics is no longer limited to comprehensive use, but more attention is paid to its surface pattern design. Affected by the diversity of ceramic shapes, the composition thickness of each part needs to be considered in the design process of graphic elements, so its design is difficult and needs to rely on intelligent layout algorithms for layout. Conventional graph element layout algorithm has poor layout effect, which does not meet the current ceramic design requirements. Therefore, this paper designs a new intelligent layout algorithm for graph elements in ceramic design based on

Table 3.3: Experimental Results

Thickness	The graph element distribution equilibrium index of the intelligent layout method for ceramic design drawings based on genetic algorithms designed in this paper	Distribution Equilibrium Index of Elements in Conventional Intelligent Layout Methods for Ceramic Design Drawings
0.2E-03	0.9611	0.4513
0.06E-03	0.9412	0.5362
0.06E-02	0.9398	0.4844
0.03E-03	0.9445	0.5512
0.2E-02	0.9512	0.4541
0.6E-03	0.9453	0.4696
0.06E-03	0.9365	0.5314
0.2E-03	0.9858	0.4845
0.13E-03	0.9144	0.4285
0.5E-03	0.9474	0.4365
4E-02	0.9545	0.5741
0.06E-03	0.9562	0.4896

genetic algorithm. The experiment results show that the designed intelligent layout algorithm of ceramic design drawing elements based on genetic algorithm has good balance index, good layout performance, reliability, and certain application value, and has made certain contributions to improving the aesthetic quality of ceramic design drawing elements. The research's focus on a specific P2P experimental platform may limit the generalizability of the findings. Different platforms or environments may have distinct characteristics and requirements, and the effectiveness of the intelligent layout algorithm and collaborative features may vary across platforms.

## REFERENCES

- [1] M. A. AL-OBAYDI, A. RUIZ-GARCÍA, G. HASSAN, J.-P. LI, C. KARA-ZAÏTRI, I. NUEZ, AND I. M. MUJTABA, *Model based simulation and genetic algorithm based optimisation of spiral wound membrane ro process for improved dimethylphenol rejection from wastewater*, Membranes, 11 (2021), p. 595.
- [2] M. ALSHEHRI, P. SHARMA, R. SHARMA, AND O. ALFARRAJ, *Motion-based activities monitoring through biometric sensors using genetic algorithm.*, Computers, Materials & Continua, 66 (2021).
- [3] W. BATAYNEH, A. BATAINEH, AND M. JARADAT, *Intelligent adaptive fuzzy logic genetic algorithm controller for anti-lock braking system*, International Review on Modelling and Simulations, 14 (2021), pp. 44–44.
- [4] R. K. BEHERA, B. P. SAMAL, S. C. PANIGRAHI, K. MUDULI, A. MOHAMED, AND A. SAMAL, *Evaluation of magnesium recovery in al-mg alloys produced by modified stir casting method using genetic algorithm optimisation technique*, International Journal of Materials Engineering Innovation, 12 (2021), pp. 134–148.
- [5] I. BEKMEZCI, M. ERMIS, AND E. B. CIMEN, *A novel genetic algorithm-based improvement model for online communities and trust networks*, Journal of Intelligent & Fuzzy Systems, 40 (2021), pp. 1597–1608.
- [6] Z. BO, Z. XIAOSI, L. YUXIN, Y. BIN, AND Z. YANG, *Combining application of wavelet analysis and genetic algorithm in wind tunnel simulation of unidirectional natural wind field near a sand ground surface*, Review of Scientific Instruments, 92 (2021).
- [7] Y. CHEN, *Location and path optimization of green cold chain logistics based on improved genetic algorithm from the perspective of low carbon and environmental protection*, Fresenius Environmental Bulletin, 30 (2021), pp. 5961–5973.
- [8] K. DAVOUDI AND P. THULASIRAMAN, *Evolving convolutional neural network parameters through the genetic algorithm for the breast cancer classification problem*, Simulation, 97 (2021), pp. 511–527.
- [9] S. EMIROGLU AND Y. UYAROGLU, *Genetic algorithm (ga)-based delay feedback control of chaotic behavior in the voltage mode controlled direct current (dc) drive system*, Zeitschrift für Naturforschung A, 76 (2021), pp. 13–21.
- [10] D. C. HIENER AND G. R. HUTCHISON, *Pareto optimization of oligomer polarizability and dipole moment using a genetic algorithm*, The Journal of Physical Chemistry A, 126 (2022), pp. 2750–2760.
- [11] Y. JIANG, *Planning of vehicle routing with mixed time windows based on the improved genetic algorithm (iga)*, in Journal of Physics: Conference Series, vol. 1952, IOP Publishing, 2021, p. 042003.
- [12] 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.
- [13] D. LIANG AND F. YANG, *Power grid environment and inductor design for power loss optimization with boost converter by genetic algorithm*, in IOP Conference Series: Earth and Environmental Science, vol. 706, IOP Publishing, 2021, p. 012022.
  - [14] S. MAHESWARAN, S. SATHESH, G. SARAN, AND B. VIVEK, *Automated coconut tree climber*, in 2017 International Conference on Intelligent Computing and Control (I2C2), IEEE, 2017, pp. 1–6.
  - [15] S. Y. MARTOWIBOWO, B. KEMALA DAMANIK, ET AL., *Optimization of material removal rate and surface roughness of aisi 316l under dry turning process using genetic algorithm*, Manufacturing Technology, 21 (2021), pp. 373–380.
  - [16] S. MIA, M. REZA, AND M. M. RAHMAN, *Genetic algorithm and fisher discriminant analysis based wavelet thresholding for speckle noise filtering in ultrasound images*, International Journal of Computer Applications, 174 (2021), pp. 13–18.
  - [17] B. NAGY AND P. SZOKOL, *A genetic algorithm for the minimum vertex cover problem with interval-valued fitness*, Acta Polytech. Hungarica, 18 (2021), pp. 105–123.
  - [18] A. M. NASRABADI AND M. MOGHIMI, *Energy analysis and optimization of a biosensor-based microfluidic microbial fuel cell using both genetic algorithm and neural network pso*, international journal of hydrogen energy, 47 (2022), pp. 4854–4867.
  - [19] T. PAN, Y. ZHANG, F. SU, V. LYNE, F. CHENG, AND H. XIAO, *Practical efficient regional land-use planning using constrained multi-objective genetic algorithm optimization*, ISPRS International Journal of Geo-Information, 10 (2021), p. 100.
  - [20] N. PERTIWI, F. S. AKBAR, E. SETIJADI, AND G. HENDRANTORO, *Linear array thinning with cavity backed u-slot patch antenna using genetic algorithm*, Journal of Science and Applicative Technology, 5 (2021), pp. 9–16.
  - [21] T. PETROV, *Modification of the genetic algorithm for comprehensive topological optimization of the synchronous motors rotor*, Power engineering: research, equipment, technology, 23 (2021), pp. 70–79.
  - [22] Z. RAN, W. MA, C. LIU, AND J. LI, *Multi-objective optimization of the cascade parameters of a torque converter based on cfd and a genetic algorithm*, Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering, 235 (2021), pp. 2311–2323.
  - [23] A. TÜRK, D. SARAL, M. ÖZKÖK, AND E. KÖSE, *Minimization of total trim loss occurring in pipe cutting in shipbuilding with genetic algorithm*, Proceedings of the Institution of Mechanical Engineers, Part M: Journal of Engineering for the Maritime Environment, 235 (2021), pp. 695–704.
  - [24] B. VIVEK, S. MAHESWARAN, P. KEERTHANA, S. SATHESH, S. BRINGERAJ, R. A. SRI, AND S. A. SULTHANA, *Low-cost raspberry pi oscilloscope*, in 2018 international conference on intelligent computing and communication for smart World (I2C2SW), IEEE, 2018, pp. 386–390.
  - [25] J. WU, Y. JI, AND S. LI, *Genetic-algorithm-based convolutional neural network for robust time series classification with unreliable data.*, Sensors & Materials, 33 (2021).
  - [26] 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.

*Edited by:* Sathishkumar V E

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

*Received:* May 15, 2023

*Accepted:* Jul 19, 2023



## DISTRIBUTION NETWORK PLANNING METHOD CONSIDERING THE COUPLING OF TRANSPORTATION NETWORK AND DISTRIBUTION NETWORK

**Abstract.** This paper introduces a novel distribution network planning method that addresses the limitations of conventional approaches. The existing methods primarily focus on optimizing component objectives using reliability analysis, which results in inadequate operational power control performance due to neglecting the coupling degree analysis of distribution network subprojects. To overcome this limitation, the proposed method incorporates the coupling of the traffic and distribution networks into the planning process. The method involves modeling the transportation network and analyzing the coupling characteristics of the planning items. Specifically, the energy efficiency coupling degree is calculated to assess the degree of coupling. Based on this analysis, the planning nodes are strategically deployed, and a comprehensive planning model is constructed. The model is then subjected to constraints and solved to obtain an optimal distribution network planning scheme. To evaluate the effectiveness of the proposed method, experiments are conducted to assess its operational power control capability. The experimental results demonstrate that when the proposed method is employed for distribution network planning, it reduces operating power and achieves a more desirable planning outcome. The novelty of this work lies in integrating the coupling analysis of the traffic network and the distribution network into the planning process. Considering the interdependencies between these networks, the proposed method enables a more comprehensive and efficient distribution network planning scheme. This approach enhances operational power control performance and improves the overall effectiveness of distribution network planning.

**Key words:** transportation networks; distribution network coupling; planning methods; linear programming.

**1. Introduction.** The main considerations for distribution network planning are planning demand, regional traffic conditions, and local construction levels. The traffic situation in the area will affect the configuration of the distribution network nodes, while the local construction level will have a limiting effect on the effectiveness of the distribution network planning. In general, the planning demand is based on economic efficiency. Different objective functions can be constructed by setting indicators such as planning cost economy, expected production economy and energy efficiency. The volume of traffic in the regional traffic situation and the route's load capacity also impact the planned route. To facilitate the study, the mathematical planning of the distribution network generally defaults to a level of construction that meets the planning needs. Currently, distribution network route planning can be carried out mainly by relying on models or expert commentary methods.

Model construction means that parameters such as the distribution network's energy efficiency or construction costs are used as objective functions, which are constructed and solved to obtain the optimal planning solution under this result. The evaluation analysis method combines principle component analysis or grey correlation method to analyze the feasibility of the distribution network planning results. By calculating the distribution network benefits or planning costs under different planning results, the best planning results can be obtained. Both of these methods are commonly used in distribution network planning projects and can produce optimal planning solutions for different planning needs, which is a goal-oriented planning method.

However, there are some limitations to these two planning methods. Firstly, the model construction method can only be used for a single planning objective. In contrast, in planning large distribution networks, there are often multiple benefits to be considered, including economic, social, and production benefits [17]. Many large distribution networks are planned to ensure the lowest possible construction costs while requiring the highest possible operational efficiency. Model-building methods cannot meet the multiple planning needs of large distribution network projects. The expert commentary method, based on the collection of expert questionnaires and the analysis of expert opinions, is highly subjective, cannot accurately and objectively analyze the actual benefits of different planning projects, and is not comprehensive enough to meet the planning needs in terms of reliability. In addition to their limitations, the two planning methods also share a common limitation, namely the excessive coupling of sub-projects [10]. The degree of sub-project coupling refers to the overlap of energy

Table 2.1: Traffic Network Influencing Factors

Elements of influence	Specific explanations
Natural factors	Terrain
	Geological conditions
	River flow
Social factors	Urban distribution
	Economic conditions
Technological factors	Technological level
	Level of road construction

and output benefits between different sub-projects in the planning results. Both conventional planning methods do not consider the collar of sub-projects, resulting in an overlap of energy efficiency between sub-projects in the planning results. Therefore, to avoid the limitations of the above planning method, the coupling between the sub-projects needs to be analyzed, and by identifying the deeper coupling between the sub-projects and dealing with the duplication, each sub-project can achieve its energy efficiency, thus ensuring that the energy efficiency of the overall project is maximized.

By analyzing the coupling characteristics of the distribution network sub-projects, the cost of construction resources can be effectively allocated when input planning, thus reducing planning expenditure and providing higher economic efficiency. In response, this paper proposes a new distribution network planning method that incorporates transport network and distribution network coupling factors into the planning scheme, and improves the economic planning efficiency of the distribution network by calculating the coupling degree of sub-projects and constraining the planning results so that sub-projects with duplicated energy benefits can be processed [9].

Ultimately, the motivation behind this research is to contribute to advancing distribution network planning methods, provide more accurate and efficient planning schemes, and support the development of sustainable and optimized distribution networks. By addressing the limitations of existing methods, the research aims to offer practical solutions to enhance the operational performance of distribution networks and meet the evolving demands of modern power systems.

**2. Traffic network model construction and planning project benefit coupling degree calculation.** To carry out rational planning of the distribution network, this paper first models the traffic network and analyses the coupling situation of the planned projects based on which the grid planning nodes are deployed. Considering that the traffic flow has a large degree of influence on the project site planning, the traffic network around the distribution network is firstly modeled [1]. The traffic network mainly consists of passing road sections and traffic line hub nodes, and one passing road section will exist in every two traffic hub nodes. Each traffic section corresponds to its own traffic flow and load capacity, determining the maximum travel time of different traffic sections. In addition, there are combinations of traffic demand nodes in the area around the distribution network, each of which includes a start node and an end node for the realization of different sections of traffic demand [4]. Depending on the traffic demand, the node combinations can generate different traffic routes, for example, a start node can match multiple end nodes, and the routes can be found efficiently by using the depth search algorithm. The main influencing elements of the traffic network are shown in Table 2.1.

Table 2.1 shows that transportation networks are affected by three main factors: natural, social, and technological. Since the issue discussed in this paper is distribution network planning, this paper focuses on social factors to analyze it, i.e., economic conditions and traffic demand factors [3]. In this regard, the traffic flow on each driving route will be accumulated on different traffic routes, and the traffic flow accumulation formula is shown below.

$$x_a = \sum_w \sum_k f_k^w \delta, \forall w, \forall k \in T_k^w, \forall a \in T_A \quad (2.1)$$

where  $x_a$  represents the accumulated traffic flow of road section  $a$  in a certain statistical period,  $w$  represents

a traffic demand node combination,  $k$  represents the driving route corresponding to the node combination;  $f_k^w$  represents the traffic flow under the driving route corresponding to the traffic demand node combination;  $\delta$  represents the accumulated times,  $T_k^w$  represents the set of all driving routes under the traffic demand node combination,  $T_A$  represents the set of passing road sections [20]. At the same time, the traffic volume under each combination of traffic demand nodes should meet the route demand, so the following expression can be obtained.

$$\sum_{k \in T_k^w} f_k^w = q_w \tag{2.2}$$

where,  $q_w$  represents the traffic flow demand under the route for a certain combination of traffic demand nodes in a certain statistical period. In this paper, the road section travel time is chosen as the main indicator to measure the traffic capacity of a traffic section, and the specific calculation formula is shown below.

$$t_a = t_a^0 \left[ 1 + 0.15 \left( \frac{x_a}{c_a} \right)^2 \right] \tag{2.3}$$

where  $t_a$  represents the passage time of the road section a,  $c_a$  represents the maximum traffic flow under the passage section,  $t_a^0$  represents the free passage time under the passage section [11]. Through the above formula can be seen that when the traffic flow of section a is lower than the maximum traffic flow, the team is in a free flow state; when the traffic flow of area a exceeds the maximum traffic flow, at this moment, can be seen as a traffic congestion situation, the passage time will be affected by the actual traffic flow [11].

After completing the above analysis for the traffic network, the distribution network benefit coupling needs to be analyzed. In the distribution network planning, there is benefit coupling between different planning sub-projects [19]. In other words, the functions of two sub-projects in a distribution network planning project may be duplicated, thus resulting in a less efficient operation of the distribution network as a whole [5]. A reasonable analysis of the coupling of distribution network benefits can effectively ensure that the sub-projects do not have overlapping parts in terms of functions and benefits, and ensure the maximum efficiency of the distribution network project. Firstly, assuming that the distribution network planning projects are  $x_i$ ,  $I$  represents the total number of expected planning projects, and assuming that the planning timing expectation is  $\zeta_i$ , the timing benefit function of distribution network benefit coupling can be constructed, and the specific function expression is shown below [6].

$$f(\zeta_i) = \begin{cases} \frac{G_i - G_{\min}}{G_{\max} - G_{\min}}, \zeta_i \in + \\ \frac{G_{\max} - G_i}{G_{\max} - G_{\min}}, \zeta_i \in - \\ 0 \leq f(\zeta_i) \leq 1 \end{cases} \tag{2.4}$$

where  $G_i$  represents the efficacy parameter generated after the distribution network planning is completed;  $f(\zeta_i)$  represents the orderliness of the planning result, the value of which ranges from 0 to 1, and the closer the value is to integer, the higher the orderliness of the planning result at this moment [18].  $G_{\max}$  and  $G_{\min}$  represent the steady-state values of the planning items in the case of maximum and minimum planning orderliness, respectively. According to the above equation, the timing function of the coupling benefits of the distribution network can be obtained, and the above timing expectation values are weighted to obtain the following expressions.

$$\zeta = [\zeta_1, \zeta_2, \dots, \zeta_I] \cdot \begin{bmatrix} \omega_1 \\ \omega_2 \\ \dots \\ \omega_I \end{bmatrix} \tag{2.5}$$

where represents the coupling weight value of the  $i$ th subproject for the  $j$ th subproject in the distribution network project. The coupling benefit equation for the distribution network project planning can be constructed by

Table 3.1: Correspondence between the main station of the distribution network and the regional scale

Regional scale	Distribution network main station size
Large cities	Large Main Station
Medium-sized cities	Medium Main Station
Small Cities	Small main station
County	Front extension main station

combining the above two formulas, and the specific expression is shown below [12].

$$O = \frac{\prod_{i=1}^I \zeta_i}{\left[ \prod_{i=1,2,\dots,I-1} (\zeta_i + \zeta_{i+1}) \cdot (\zeta_i + \zeta_{i-1}) \right]^{1/I}} \quad (2.6)$$

The value of  $O$  represents the benefit coupling degree, which ranges from 0 to 1. The closer the value is to integer, the higher the efficacy overlap between the sub-projects in the planning result at this moment. The above calculation of project effectiveness coupling degree can effectively correct the planning results [7].

**3. Distribution network project planning node deployment.** The planning of the distribution network project needs to be reasonably matched with the regional construction scale and the total number of sites, as shown in Table 3.1.

According to the above table, it can be seen that the corresponding site size varies for different size cities. For larger cities, a large master station needs to be matched, while for smaller counties and other areas, the distribution grid master station can be selected from the front extension master station [17]. In this paper, we focus on large cities with corresponding large distribution network master stations for distribution network planning analysis. In general, the reliability level of the distribution network master station in large cities is higher, and the distribution equipment level is higher, and the power supply mode of composite power supply is generally adopted so that the power can be restored faster for some power outages or blackouts, which also puts more requirements on the distribution network planning of large cities. First of all, the distribution network needs to be able to supply the required amount of power with minimum energy consumption. Therefore, when planning the distribution network, it is necessary to take into account the demand for power in different areas and the distance from the proposed planning center [15]. For areas with high demand, the distance between the distribution network center and the area needs to be reduced in order to reduce transmission losses.

The conventional distribution network planning method is mainly through the analysis of the planning demand, based on which the planning nodes are deployed in combination with the power supply efficiency of each region, and then different nodes are connected between them, and the planning path can be determined. The planning results obtained by this method have more planning routes, which can be reasonably selected for different planning needs and the specific geological conditions of the study area. However, due to the lack of consideration of the planning cost, the planning results under the traditional method have low planning efficiency and the algorithm has low robustness and cannot complete the corresponding path search task within the specified time. Therefore, in this paper, the conventional distribution network planning method is optimized, and the optimal planning path is obtained by searching the planning paths under different nodes with the planning efficiency as the objective function, so that the coupling efficiency of the planning paths can be effectively avoided [8].

In this regard, firstly, the nodes of the planning project need to be reasonably deployed, and the specific nodes need to be spiritually adjusted according to the power supply demand of the study area and the construction level of electrical energy equipment. In this paper, six planning nodes are selected for the power supply demand characteristics of the main station of large urban distribution network, and the construction level of distribution network under each node can meet its corresponding power supply demand. The nodes are barrier-free and can be connected flexibly. First of all, the planning paths connected by the distribution network planning nodes need to have a single nature. Too many repetitive planning paths not only cause waste



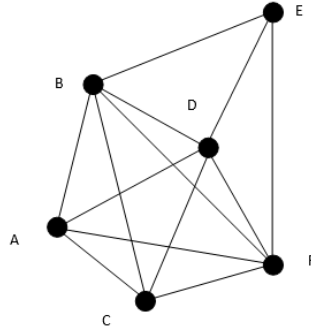


Fig. 3.1: Initial planning results of distribution grid gridding

of planning resources, but also create a large electric pressure on the operation of the main distribution network in large cities. Therefore, it is necessary to search for the best planning paths under different nodes in order to maximize the benefits of the planning results. In order to meet the planning requirements of the distribution network master in large cities, the planning paths under each node need to be analyzed to avoid duplicate paths as much as possible [2].

In order to improve the planning effect, this paper chooses the grid method to divide the planning nodes of the project, for which, the gridization parameters need to be simulated. Combined with the algorithm, it can be assumed that the initial set of ant colony size is  $V$ , and the initial set of planning lines is  $E$ . The initial planning results can be obtained through the ant colony seeking control, as shown in Figure 3.1.

In the initial planning result of the above figure, assuming that the total length of the distribution grid gridding search path is  $p_{ij}$ , the grid connection matrix expression can be obtained as shown below.

$$A = p_{ij} \times e^t \quad (3.1)$$

where  $e^t$  represents the search path parameter. Since each individual in the ant colony population has variability in search performance, individual variability is also taken into consideration when grid planning is performed for the distribution network, and the obtained grid node planning degree is calculated as shown below.

$$d_i = \frac{\sum_{i=1}^V V \times D}{r_i} \quad (3.2)$$

where  $D$  represents the diagonal element due to the distribution grid gridding, the value is mainly related to the joint probability density, and  $r_i$  represents the characteristic value of the difference in search performance of different individuals under the ant colony population. Assuming that the distribution grid gridding chunking matrix is, the total number of ant colony collections under the optimal planning route can be calculated according to the above formula, which is shown below.

$$U = \frac{\ln(V \times E)}{u_{ij}} \quad (3.3)$$

where, represents the probability of effective connection distribution, through the above steps can be completed for the initialization of the grid results, the final results of the optimization of node deployment, as shown in Figure 3.2.

The above figure shows that, compared with the initial grid planning results, the optimized node deployment structure diagram in this paper has only one path between each node and it is the optimal path, so it can satisfy the power supply efficiency while controlling the energy consumption to the minimum, thus improving the energy efficiency of distribution network planning [14]. The above steps can be completed for the effective deployment of the planning nodes to ensure a single node planning path, which helps to provide comprehensive planning benefits.

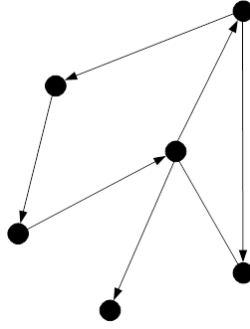


Fig. 3.2: Node deployment structure under optimal path planning

**4. Distribution network planning model construction.** After the construction of the traffic network model and the calculation of the coupled benefit values, the distribution network planning model is constructed by combining the grid diagram of the planning nodes deployed in this paper with the energy efficiency as the objective function and the linear regression method. The control expression of energy efficiency of distribution network is shown as follows [16].

$$\max g(x) = mk_1x_1 + k_2x_2 + \dots, k_nx_n \quad (4.1)$$

where  $g(x)$  represents the standardized control model of energy efficiency of the distribution network,  $k$  represents the efficiency parameter of the planning project,  $m$  represents the linear synchronization parameter, and  $n$  represents the number of linear planning. Considering the energy cost and expected energy efficiency of the distribution network construction, this paper takes the maximum energy efficiency constraint as the objective function, and the obtained function expressions are shown below.

$$\max \sum C(P_{en,i}) = \sum C_1(P_{DG,L}) + \sum C_k(P_{B,H}) \quad (4.2)$$

where  $P_{en,i}$  represents the maximum energy efficiency under the  $i$ th planning path under linear planning;  $P_{DG,L}$  represents the energy efficiency load under the standardized energy efficiency control model;  $P_{B,H}$  represents the optimal control objective under the source network load;  $C_k$  represents the energy load parameters of the distribution network;  $C_1$  represents the energy cost consumed by the distribution network construction,  $C$  represents the total energy consumption of the planning path,  $B$  represents the energy efficiency control objective;  $H$  represents the constraint parameters of the objective function. On the basis of the above objective function, the constraints are applied to the planning model. First, the energy consumption brought by the operating load of the distribution network under different planning results needs to be calculated as shown in the following formula [13].

$$P'_{F,r} = P_{F,r} + P_{loss,r} \quad (4.3)$$

where  $P'_{F,r}$  represents the total power consumption values of electrical equipment due to operating load in the distribution network,  $P_{F,r}$  represents the power consumption of individual electrical equipment in the normal state,  $P_{loss,r}$  represents the power consumption of the distribution network in a single path under the optimal planning line,  $F$  represents the energy load of the distribution network,  $r$  represents the storage load, and  $loss$  represents the transmission path with the largest energy consumption. In general, the energy consumption generated by the main station of the distribution network in a large city should run the load in the operation process is generally within 10%, so this paper takes the historical operation data of a distribution network as the research object and constructs the constraints of the objective function, i.e., the upper and lower limits of the operation energy consumption, by analyzing the operation energy consumption ratio in the historical data, as shown in Table 4.1.

Table 4.1: Constraints

Control items	Upper limit conditions	Lower limit conditions
$P_{DG,L}$	$P_{DG,L}(\max) \leq L$	$P_{DG,L}(\min) \geq H$
$P_{B,H}$	$P_{B,H}(\max) \leq H$	$P_{B,H}(\min) \geq B$
$P_{F,r}$	$P_{F,r}(\max) \leq r$	$P_{F,r}(\min) \geq L$

Based on the above constraints, the objective function is solved to achieve a reasonable planning for the distribution network. The specific solution process is shown below.

Firstly, the optimal control boundary of the energy efficiency of the distribution network is calculated, and the value is mainly obtained by summing the upper limit condition parameters and the lower limit condition parameters of the control project. Then the energy efficiency control model is calculated for the specific model parameters of the energy transmission power, and the actual increment is calculated according to the size of the transmission power, and the specific calculation formula is shown below.

$$\Delta P = P_{T0} - P_T \quad (4.4)$$

where  $\Delta P$  represents the actual incremental power output of the distribution network,  $P_{T0}$  represents the statistical output power value, and  $P_T$  represents the warning output power under the energy efficiency control model. The above equation can control the actual incremental deviation of output power, and select a suitable planning path according to the ideal deviation value, and then optimize the planning path to ensure the optimal energy efficiency of the distribution network under the planning path and to meet the regional power supply demand. Through the above steps, the distribution network planning model can be constructed, and the control planning effect of the model is constrained by combining the constraints designed in this paper, and the model is solved to obtain the best planning path and scheme. Combining the contents of this section with the above-mentioned results of traffic network model construction and coupling benefit calculation, and grid-based planning node deployment, the design of the distribution network planning method considering the coupling of traffic network and distribution network is completed. The proposed method aims to control the actual incremental deviation of output power in order to optimize the planning path and ensure the optimal energy efficiency of the distribution network while meeting the regional power supply demand. By selecting a suitable planning path based on the desired deviation value, the planning path is further optimized to achieve the desired control effect. The construction of the distribution network planning model involves combining the formulated constraints and utilizing the designed control measures outlined in this paper. By solving the model, the best planning path and scheme can be obtained, ensuring efficient and reliable operation of the distribution network. Furthermore, the completed design of the distribution network planning method takes into account the coupling of the traffic network and the distribution network. This includes the construction of the traffic network model, calculation of coupling benefits, and deployment of planning nodes based on the grid. By integrating these components with the aforementioned steps, a comprehensive and effective distribution network planning method is achieved.

## 5. Experiment and analysis.

**5.1. Experimental preparation.** In order to prove that the distribution network planning method considering the coupling of traffic network and distribution network proposed in this paper is better than the conventional distribution network planning method in terms of planning classification effect, after the design of the theoretical part is completed, an experimental session is constructed to test the actual planning effect of the method in this paper. In order to ensure the experimental effect, two conventional distribution network planning methods are selected for comparison, namely the distribution network planning method based on reliability analysis and the distribution network planning method based on temporal correlation.

The experimental object selected for this experiment is a 24-node distribution network system with a rated voltage of 20 kV under the system, which contains a total of four main distribution stations and 10 power supply demand nodes, the specific structure of which is shown in Figure 5.1.

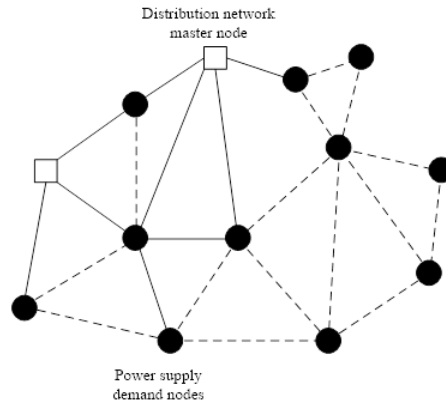


Fig. 5.1: Distribution network system node diagram

Table 5.1: Distribution network modeling parameters

Modelling parameters	Specific configuration
Power Plate Cap/MV	55
Lower Power Plate Limit/MV	60
O&M factor	28.63
Installation cost	30.1
Source supply climb rate/M	150
Source Network Supply Down Climb Rate	165

In order to improve the accuracy of the experimental results, MATLAB software is used to model the above distribution network system in this experiment. To facilitate the modeling, line parameters of the original distribution network system and substation-related data are extracted in this paper to facilitate the subsequent rational planning of the distribution network for energy efficiency, and the specific modeling parameters are shown in Table 5.1.

The distribution network system was modeled according to the contents of Table 5.1, and the modeling results were planned using three methods to compare the energy efficiency costs under different planning methods.

**5.2. Analysis of test results.** The comparison criterion chosen for this experiment is the planning performance of the planning method, and the specific measurement index is the optimal control power of the distribution network under different planning methods, the lower the value represents the better planning effect of the algorithm, and the specific experimental results are shown in Figure 5.2.

The above experimental results show that the optimal control power of the distribution network model under different planning methods also varies at different control times. By observing the power change curves, it can be seen that the control power of the distribution network under the two conventional planning methods is higher, which indicates that the energy consumption of the distribution network is larger at this moment. The model with the coupling of traffic network and distribution network proposed in this paper has lower optimal control power, which proves that the planning method proposed in this paper has better planning performance and can reasonably control the operating power of the distribution network while meeting the planning requirements.

**6. Conclusion.** This research paper proposes a distribution network planning method that considers the coupling of the traffic network and the distribution network. By incorporating the traffic network model and analyzing the coupling degree of sub-projects, the distribution network planning model is constructed to

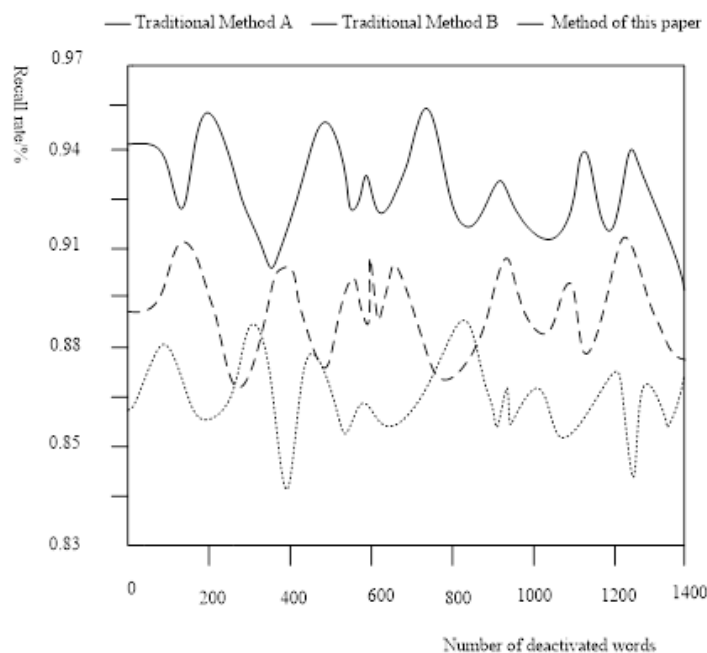


Fig. 5.2: Algorithm recall comparison results

effectively control the operation of the power grid. This approach ensures the distribution network meets the power supply demand in the area while minimizing power wastage, resulting in improved economic and energy efficiency.

Future research should focus on investigating the distinct power characteristics of AC and DC systems in greater detail. Understanding these differences will contribute to further enhancing the effectiveness and applicability of the distribution network planning method. By exploring the unique requirements and challenges posed by AC and DC systems, it will be possible to develop more tailored and optimized planning strategies for both types of power networks. This will ultimately contribute to the advancement of energy-efficient and sustainable distribution network planning practices.

#### REFERENCES

- [1] I. C. BARUTCU, *Examination of the chance constrained optimal wt penetration level in distorted distribution network with wind speed and load uncertainties*, Engineering, Technology & Applied Science Research, 11 (2021), pp. 7311–7320.
- [2] P. BHATT, *Harmonics mitigated multi-objective energy optimization in pv integrated rural distribution network using modified tlbo algorithm*, Renewable Energy Focus, 40 (2022), pp. 13–22.
- [3] B. CANIZES, V. SILVEIRA, AND Z. VALE, *Demand response and dispatchable generation as ancillary services to support the low voltage distribution network operation*, Energy Reports, 8 (2022), pp. 7–15.
- [4] G. S. CHAWDA AND P. KUMAR, *Mitigation of nonlinear load influence on the power quality of the low-voltage distribution network using dstatcom*, in Active Electrical Distribution Network, Elsevier, 2022, pp. 229–241.
- [5] R. CHENG, Q. LI, Y. ZHANG, W. DENG, Y. AN, AND M. CHEN, *Three-phase power cooperative management of distribution network with high distributed energy resources penetration*, Energy Reports, 8 (2022), pp. 239–250.
- [6] C. DING, T. NIE, AND Y. CHEN, *A distribution network solid-state dc circuit breaker with current limiting function*, Energy Reports, 8 (2022), pp. 986–994.
- [7] Y. DING, T. LU, Z. WANG, H. WU, AND X. LU, *A full time scale resilience improvement strategy of distribution network under extreme weather*, Energy Reports, 8 (2022), pp. 1019–1025.
- [8] W. GUO AND H. WANG, *A protection scheme for active distribution network based on the cooperative control of inverter-interfaced distributed generators*, Energy Reports, 8 (2022), pp. 342–352.

- [9] L. GUTIERREZ-LAGOS, K. PETROU, AND L. OCHOA, *Quantifying the effects of mv-lv distribution network constraints and der reactive power capabilities on aggregators*, IET Gener. Transm. Distrib, 15 (2021), pp. 2019–2032.
- [10] I. M. JAWAD AND W. S. MAJEED, *Congestion mitigation in distribution network by integrated distributed generations for improving voltage profiles and minimizing the losses*, Journal of Engineering and Sustainable Development, 25 (2021), pp. 78–87.
- [11] B. KHAN, K. REDAE, E. GIDEY, O. P. MAHELA, I. B. TAHA, AND M. G. HUSSIEN, *Optimal integration of dstatcom using improved bacterial search algorithm for distribution network optimization*, Alexandria Engineering Journal, 61 (2022), pp. 5539–5555.
- [12] W. LIN, B. XIA, Y. TANG, T. ZHANG, AND Y. QI, *Aoi research on pmu cloud side cooperative system of active distribution network*, Energy Reports, 8 (2022), pp. 677–686.
- [13] Y.-J. LIN, C. JING, C. CAO, AND K. BAI, *Distributed power supply access distribution network planning with timing characteristics*, Computer Simulation, 39 (2022), pp. 51–55.
- [14] P. J. MITCHELL, S. G. BOLAM, H. L. CLOSE, C. GARCIA, J. MONK, AND K. ALLIJI, *Distribution of demersal fish assemblages along the west coast of st lucia: Implications for planning no-take marine reserves*, Aquatic Conservation: Marine and Freshwater Ecosystems, 31 (2021), pp. 1354–1366.
- [15] H. NYKYFORCHYN, L. UNIGOVSKYI, O. ZVIRKO, M. HREDIL, H. KRECHKOVSKA, O. STUDENT, AND O. TSYRULNYK, *Susceptibility of carbon pipeline steels operated in natural gas distribution network to hydrogen-induced cracking*, Procedia Structural Integrity, 36 (2022), pp. 306–312.
- [16] W. SONG AND J. ZHU, *A goal-reference-point decision-making method based on normal cloud model and its application in distribution network planning evaluation*, Information Sciences, 577 (2021), pp. 883–898.
- [17] Q. WEI, J. ZHENG, W. TANG, AND Q. WU, *Equivalent model of active distribution network considering uncertainties of wind turbines, photovoltaics and loads*, in IOP Conference Series: Earth and Environmental Science, vol. 645, IOP Publishing, 2021, p. 012086.
- [18] Z. WU, L. YE, P. LIAO, C. LI, Z. ZHOU, S. LI, AND J. SUN, *State estimation of distribution network based on hybrid measurement combined with multi-source asynchronous data*, Energy Reports, 8 (2022), pp. 1778–1783.
- [19] R. XU, C. CHEN, D. CHEN, G. ZHOU, Y. JIN, X. WU, W. WANG, X. FENG, S. LIU, AND Z. LIN, *Maximum openable capacity optimization method of active distribution network considering multiple users access*, Energy Reports, 8 (2022), pp. 43–50.
- [20] K. YU, B. XUE, F. GU, H. HUA, Y. YUAN, AND Q. LI, *A novel pet model based volt/var control of ac-dc hybrid distribution network*, Energy Reports, 8 (2022), pp. 2672–2685.

*Edited by:* Sathishkumar V E

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

*Received:* May 15, 2023

*Accepted:* Jul 9, 2023



## REMOTE MONITORING SYSTEM OF DIGITAL AGRICULTURAL GREENHOUSE BASED ON INTERNET OF THINGS

**Abstract.** In the actual operation process of the conventional digital agricultural greenhouse monitoring system, there are problems such as limited monitoring scope and large deviation between the monitoring results and the actual situation of the greenhouse. To solve this problem, a new remote monitoring system is proposed by introducing the technology of the Internet of Things. On the basis of the completion of the hardware design of the remote monitoring system, the optimal fusion data value of the remote monitoring of the digital agricultural greenhouse is obtained by establishing the monitoring data fusion model. The particle swarm optimization fuzzy control algorithm is designed to optimize the adaptive remote monitoring process of the system dynamically. The Internet of Things technology is used to deploy the remote monitoring system of digital agricultural greenhouses online to fully ensure the quality and timeliness of the remote monitoring system. The test results show that the new system can significantly improve the greenhouse remote monitoring deviation, and the monitoring value is close to the actual value.

**Key words:** Internet of Things; System; Digitization; Agricultural greenhouses; Monitor; Long-range

**1. Introduction.** With the continuous development of greenhouse planting industry, its planting scale is also expanding, and the greenhouse has gradually achieved the goal of digital development. Relying on the Internet of Things and sensor technology, the management level of the greenhouse has been significantly improved [14]. Digital agricultural greenhouse refers to the installation of meteorological stations and soil moisture stations in the greenhouse to monitor the greenhouse air temperature, humidity, carbon dioxide concentration, light, soil pH, soil temperature and moisture, and other factors in real time [6]. According to the monitoring data, the management personnel can master and analyze the environment in the shed, control the opening and closing of heating and lighting, top and side windows, water and fertilizer integration equipment in the hut through a wireless network, and adjust the environment in the shed by controlling the temperature and humidity, so that the crops are always in a suitable growth environment [9]. Digital agricultural greenhouse gas (GHG) monitoring uses digital technologies, such as sensors, IoT devices, and data analytics, to measure, analyze, and manage GHG emissions in agricultural settings, particularly in greenhouse environments.

The advent of digital greenhouses has significantly lowered the labor costs associated with greenhouse farming, boosted the harvests of crops grown in such controlled environments, and increased the economic viability of greenhouse agriculture [16]. In the operation process of the digital greenhouse, a scientific remote monitoring system is needed to grasp the dynamic change information of each area in the greenhouse in real-time and adjust the temperature, humidity, light intensity, and soil pH [15] in the greenhouse in time. At present, the traditional greenhouse remote monitoring system in the actual application process is still not perfect; the monitoring range is limited, affected by interference factors, the monitoring results and the actual situation in the greenhouse have a significant deviation, and can not obtain more accurate monitoring data, and the monitoring time efficiency is poor [5]. As related disciplines constantly advance, the technology behind the Internet of Things has made significant strides forward. It has a large amount of information to obtain the overall processing capacity. Collaborative computer technology has pushed China's agriculture into a more important technological era [8]. With the Internet of Things, we can obtain a certain parameter of the digital agricultural greenhouse, send it through the sensor, conduct a comprehensive analysis of these data, and design various responses to these data to ensure that the operation of the agricultural greenhouse is controlled in a scientific way [10]. It is precisely because of the advanced management of intelligent agriculture that a large number of labor forces have been liberated, a large number of materials have been saved, and at the same time, crops have been ensured in the best production environment. Based on this, this paper introduces the Internet of Things technology and proposes a new remote monitoring system for digital agricultural greenhouses.

In general, sensors and IoT (Internet of Things) technologies can be used to monitor the emission of

greenhouse gases (GHGs). Several types of sensors are commonly employed for this purpose, including:

1. **Gas Sensors:** These sensors are designed to detect specific gases, such as carbon dioxide (CO<sub>2</sub>), methane (CH<sub>4</sub>), and nitrous oxide (N<sub>2</sub>O). They can measure the concentration levels of these gases in the atmosphere, providing valuable data for GHG monitoring.
2. **Optical Sensors:** Optical sensors utilize light absorption or scattering properties to measure GHG concentrations. For example, infrared (IR) sensors can detect and quantify the levels of GHGs based on their unique absorption patterns in specific wavelengths.
3. **Air Quality Sensors:** These sensors can monitor multiple environmental parameters, including temperature, humidity, and particulate matter. They can indirectly provide insights into GHG emissions by analyzing air quality changes associated with combustion or industrial processes.
4. **Remote Sensing:** Remote sensing techniques involve using satellite-based or aerial sensors to capture and analyze data on GHG emissions. These methods provide a broader spatial coverage and can monitor GHG sources across large areas.

The effects of greenhouse gases in agricultural applications are significant. Increased levels of GHGs in the atmosphere, such as CO<sub>2</sub> and CH<sub>4</sub>, contribute to the greenhouse effect, trapping heat and leading to global warming. This warming trend affects various aspects of agriculture:

1. **Crop Productivity:** Rising temperatures and altered precipitation patterns can impact crop growth, yield, and quality. Some crops may suffer from heat stress, reduced water availability, or increased susceptibility to pests and diseases.
2. **Water Resources:** Changes in temperature and precipitation patterns can affect water availability for irrigation and crop growth. Droughts, floods, and altered hydrological cycles can disrupt agricultural practices and water management.
3. **Soil Health:** Increased temperatures and changes in precipitation can impact soil moisture levels, nutrient availability, and microbial activity. These changes can affect soil fertility, nutrient cycling, and overall soil health, influencing crop productivity.
4. **Pest and Disease Dynamics:** Climate change can alter the geographical distribution and abundance of pests and diseases. Warmer temperatures can extend the growing season for certain pests, leading to increased pest pressure and the need for additional pest management strategies.
5. **Livestock Production:** Heat stress due to higher temperatures affects livestock health and productivity. Extreme weather events can also disrupt animal husbandry practices and feed availability.

Understanding the effects of GHGs in agricultural applications is crucial for sustainable farming practices, resource management, and adaptation strategies. By monitoring GHG emissions and implementing mitigation measures, the farm sector can work towards reducing its carbon footprint and promoting climate-resilient practices.

**2. Hardware design of digital agricultural greenhouse remote monitoring system.** In the remote monitoring system of the digital agricultural greenhouse, the data gateway is the transmission bridge of primary environmental data in the greenhouse, which is mainly composed of a microprocessor module, WiFi communication module, 485 serial communication module, and power module. The hardware structure of the remote monitoring system data gateway designed in this paper is shown in Figure 2.1.

As shown in Figure 2.1, the data gateway hardware of the digital agricultural greenhouse remote monitoring system designed in this paper takes the CPU as the core part to coordinate the processing of sensor equipment data. The 485 serial communication module is responsible for uploading the field data of the agricultural greenhouse to the microprocessor and sending the relevant control instructions [2] to the controller. The WiFi module acts as an intermediary between the gateway and cloud server, facilitating the transmission of processed data to the cloud server, receiving control instructions from the same server, and performing decoding analysis [19]. Secondly, the system microprocessor is selected and designed. According to the field investigation of the digital agricultural greenhouse, and after careful consideration of the cost, data processing speed, safety, and reliability of the system function, this paper finally selects the STM32F103 series and uses the STM32F103TRPD8 microprocessor. Set the microprocessor performance parameters, as shown in Table 2.1.

As shown in Table 2.1, the microprocessor performance parameters designed in this paper can stabilize and adjust the operating power frequency of the system through the microprocessor, reduce the interference impact



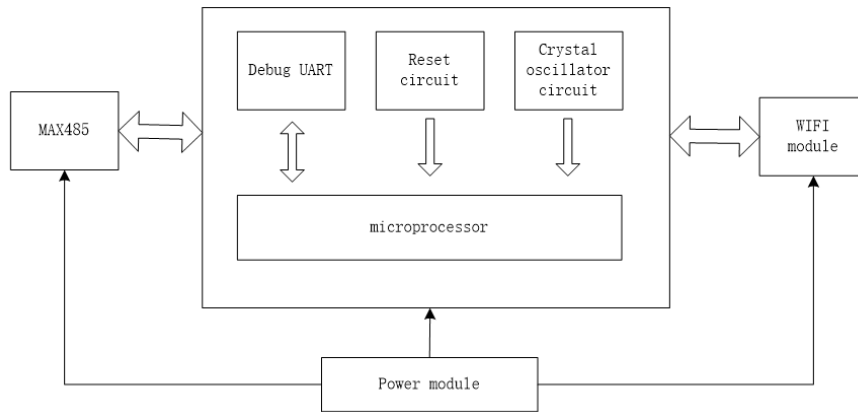


Fig. 2.1: Hardware Structure of Remote Monitoring System Data Gateway

Table 2.1: System Microprocessor Performance Parameter Settings

No	Internal structure	Functional characteristics
1	Kernel	The 32-bit Cortex M3 microcontroller core has a maximum operating frequency of 72 MHz.
2	Storage	Up to 512K bytes of flash memory, 64K bytes of SRAM.
3	Clock, reset, and power management	The I/O pin and supply voltage are 2.0-3.6V. 4-16MHz crystal oscillator embedded A 40 kHz RC oscillator and a 32 kHz RTC oscillator.
4	Low power consumption	It has three modes: sleep, shutdown, and standby, with VBAT as RTC and backup storage power supply.
5	I/O port	Up to 112 fast I/O ports, so the ports can image 16 external interrupts.
6	Communication interface	Up to 13 communication interfaces, 2 IC interfaces, 5 USART interfaces, 3 SPI interfaces, CAN interfaces, and USB 2.0 full-speed interfaces.
7	Debug Mode	With SWD and JTAG interfaces, embedded ETM tracking module.
8	A/D conversion	There are three 12-bit AD converters, 0-3.6V conversion range, triple sampling, and function retention.

caused by power fluctuations, and improve the reliability of the system circuit [13].

Design of air temperature and humidity sensor for digital agricultural greenhouse. This paper selects the RS-BYH-M air temperature and humidity sensor that Shandong Jianda Renke Company designed. The sensor is integrated with a temperature and humidity measurement structure, and its output signal is RS485. The internal power supply, induction probe, and signal output are all isolated. The probe is waterproof and sealed at the probe position, which has good waterproof and sealing performance and fully meets the requirements of agricultural greenhouse environmental monitoring [11]. Basic parameter settings of air temperature and humidity sensor are shown in Table 2.2.

Set the air temperature and humidity sensor according to the basic parameters shown in Table 2.2, so that it can accurately monitor and collect the temperature and humidity of the digital agricultural greenhouse and provide data support for remote monitoring of digital agricultural greenhouse [12].

The light intensity sensor adopts the ZZ-LRS-LIGHT light intensity sensor designed by Shandong Jianda

Table 2.2: Basic Parameters of Air Temperature and Humidity Sensor

No	Internal structure	Parameter
1	DC power supply	12V-24VDC
2	Maximum power consumption	0.4W
3	Output signal	RS485
4	Corresponding time	< 15S (1m/s wind speed)
5	Temperature long-term stability	$\leq 0.1^\circ \text{C/y}$
6	Humidity long-term stability	$\leq 1\%y$
7	Temperature measurement range	-40 $^\circ\text{C}$ -80 $^\circ\text{C}$
8	Humidity measurement range	0-100%RH
9	Temperature resolution	0.1 $^\circ\text{C}$
10	Humidity resolution	0.15RH
11	Working environment pressure range	0.9-1.1atm

Table 2.3: Parameter Settings of Soil Environment Sensor

No	Internal structure	Parameter
1	Moisture measurement range	0-100%
2	Moisture accuracy	0-53%: $\pm 3\%$ ;53-100%: $\pm 5\%$
3	Temperature measurement range	-40 $^\circ\text{C}$ -80 $^\circ\text{C}$
4	Temperature measurement accuracy	$\pm 0.5^\circ \text{C}$
5	Conductivity measurement range	0-10000us/cm
6	Conductivity resolution	10us/cm
7	Port communication	RS485 Modbus

Renke Company. The sensor adopts a high-sensitivity photosensitive probe with stable signal and high precision and uses RS485 communication. It has a wide measurement range, good linearity, good waterproof performance, convenient use, convenient installation, long transmission distance, etc. [21].

The system uses the soil temperature, humidity, and salinity three-in-one sensor RS-MTUL-GTR3 designed by Shandong Jize Company. The sensor is suitable for measuring soil temperature, soil moisture, and pH values in agricultural greenhouses. It has high precision, fast response, and stable output [4]. It is not affected by soil salinity and is suitable for various soils. It can be buried in the ground for a long time, is resistant to long-term electrolysis, has strong corrosion resistance, vacuum sealed, and is entirely waterproof [24]. Use RS485 serial Modbus standard protocol, easy access to the system, and long transmission distance. The soil environment sensor parameter settings are shown in Table 2.3.

Set the parameters of the soil environment sensor according to the parameters in Table 2.3 to ensure the reliability of its operation.

System HD video surveillance camera. The HD surveillance camera uses real-time video to monitor the growth of crops planted in the greenhouse and the dynamic changes in the greenhouse area [3]. Monitor the changes in digital agricultural greenhouses more intuitively. This system selects the PKI85-PIND5A high-definition video surveillance camera produced by Shandong Konka Ning'an Company. The real-time video surveillance function can meet the need for all-weather video surveillance for digital agricultural greenhouses [20].

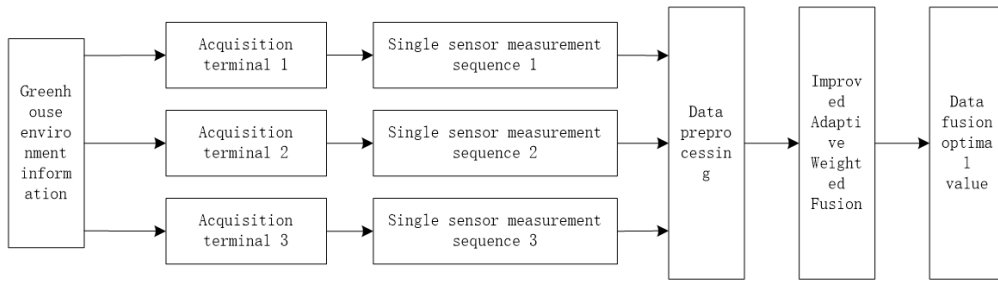


Fig. 3.1: Sensor Data Fusion Model

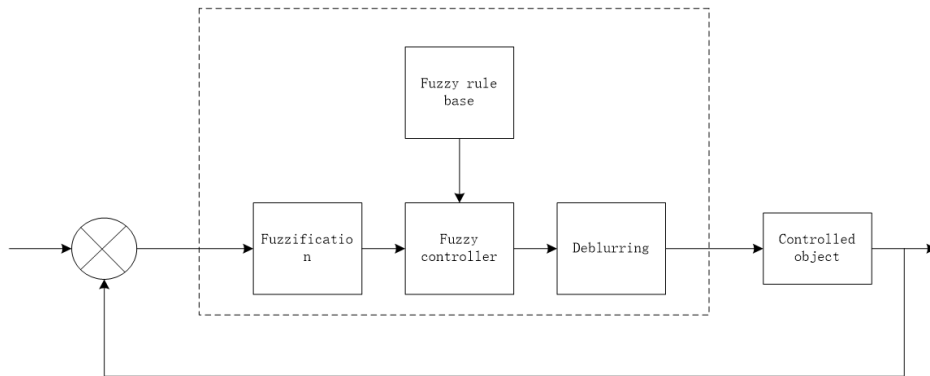


Fig. 3.2: Schematic diagram of the principle of the particle swarm optimization control algorithm

### 3. Software design of digital agricultural greenhouse remote monitoring system.

**3.1. Establish monitoring data fusion model.** The environmental data acquisition of a digital agricultural greenhouse depends on each sensor, but there will be a large deviation in the collected value due to the sensor accuracy and collection terminal failure. Based on this, this paper first conducts data preprocessing operations on the collected data of a single sensor and then fuses the collected values of similar sensors through an improved adaptive weighted fusion algorithm. The optimal fusion value [10] is obtained. The sensor data fusion model is shown in Figure 3.1.

As shown in Figure 3.1, through the iterative operation of the model, the outliers with significant measurement deviation are removed, and the optimal fusion data value of digital agricultural greenhouse remote monitoring is obtained to provide data support for subsequent remote monitoring [22].

**3.2. Design system particle swarm optimization fuzzy control algorithm.** The environment of the remote monitoring system of the digital agricultural greenhouse is complex, many factors need to be considered, and the environmental parameters interact, which makes the design of the monitoring system very challenging. The PID control algorithm and switch control method used in the monitoring system cannot cope with the multivariable and strong coupling characteristics of the greenhouse [7]. Given the limitations mentioned earlier, this paper suggests utilizing the particle swarm optimization algorithm, which can improve optimization speed and requires fewer parameter settings. Furthermore, a fuzzy PID control algorithm will also be employed for dynamic optimization. The principle diagram of the particle swarm optimization control algorithm is shown in Figure 3.2.

As shown in Figure 3.2, the specific process of fuzzy control is to collect the accurate value of the controlled object through the measuring equipment, compare it with the set value of the system, and then write the

difference into the fuzzy controller as an input [17]. First, the input value should be fuzzed through the fuzzy interface, and the measured actual value should be converted into a fuzzy vector [1]. Secondly, the fuzzy reasoning steps are carried out. The fuzzy controller operates by executing fuzzy reasoning in accordance with the fuzzy control rules outlined in the knowledge base and makes reasoning decisions on the fuzzy inputs to obtain the corresponding fuzzy output set [18]. The knowledge base consists of two parts: database and rule base. The last step in the control process is to perform the deblurring operation [23]. The blur output can be used to control or drive the actuator through the operation of the deblurring interface. With this in mind, the membership function for the system remote monitoring can be established through the implementation of the particle swarm optimization control algorithm, and the expression is:

$$\mu_e(x) = e^{-\left(\frac{x-m}{\sigma}\right)^2}$$

Among them,  $\mu_e(x)$  represents the membership function of the digital agricultural greenhouse remote monitoring system;  $m$ ,  $\sigma$  they respectively represent the system monitoring image parameters, in which,  $\sigma$  determines the width of the system monitoring function image;  $m$  determines the center point of the system monitoring function image. By designing the system particle swarm optimization control algorithm, the adaptive remote monitoring process of the system is dynamically optimized, and the overall operation performance of the system remote monitoring is improved.

The Particle Swarm Optimization Fuzzy Control algorithm (PSO-FC) is a hybrid intelligent control approach that combines Particle Swarm Optimization (PSO) and fuzzy logic to optimize and control complex systems. PSO is a population-based optimization algorithm inspired by the social behavior of bird flocking or fish schooling. Fuzzy logic, on the other hand, is a mathematical framework that deals with uncertainty and imprecise information using linguistic variables and fuzzy rules.

The PSO-FC algorithm involves the following steps:

1. Initialization: Initialize a swarm of particles, each representing a potential solution to the control problem. Each particle has a position and velocity vector.
2. Fitness Evaluation: Evaluate the fitness of each particle based on its current position in the solution space. In the context of fuzzy control, the fitness function assesses the performance of the control system based on predefined criteria or objectives.
3. Update Particle Velocity and Position: Adjust the velocity and position of each particle based on its own experience and the best experience of the swarm. PSO utilizes the concept of social learning, where particles communicate and update their positions based on their own best solution (personal best) and the best solution found by the swarm (global best).
4. Fuzzy Rule Base: Construct a fuzzy rule base that defines the relationships between system inputs, outputs, and control actions. This rule base comprises linguistic variables, fuzzy membership functions, and a set of fuzzy rules.
5. Fuzzy Inference: Apply fuzzy inference using the current system inputs and the fuzzy rule base to determine the appropriate control action or output. Fuzzy inference involves fuzzification (mapping crisp inputs to fuzzy sets), rule matching, aggregation, and defuzzification (mapping fuzzy outputs to crisp values).
6. Control Action Update: Update the control action or output based on the fuzzy inference results.
7. Termination Criterion: Repeat steps 2 to 6 until a termination criterion is met. The termination criterion can be a predefined number of iterations or when the desired control performance is achieved.

The PSO-FC algorithm combines the optimization capabilities of PSO with the reasoning capabilities of fuzzy logic to adaptively adjust control parameters and optimize control performance in complex and uncertain systems. It has been applied in various domains, including robotics, power systems, process control, and intelligent transportation systems, to achieve improved control accuracy, robustness, and adaptability.

**3.3. Online deployment based on the Internet of Things.** After completing the design of the above particle swarm optimization control algorithm, next, use the Internet of Things technology to deploy the digital agricultural greenhouse remote monitoring system online to fully ensure the quality and timeliness of the system's remote monitoring.

The system's online implementation is divided into three distinct layers - the perception layer, network layer, and application layer, facilitated by the use of Internet of Things technology. The perception layer serves as the primary data collection mechanism, capturing physical events and agricultural greenhouse data while ensuring external information is digitized. The network layer associates environmental data from the perception layer with user data available in the application layer, thereby enhancing the reliability and security of information transmission. Lastly, the application layer processes and computes large volumes of sensory information generated by the digital agricultural greenhouse while enabling seamless interaction with the internet.

Package the back-end project, use FileZilla software, upload the package file to the server-specific folder, install the project dependencies, and use the node Exe. The JS process management tool PM2 starts the project and then uses the "Pm2list" command to view the operation status after the project is started. After the deployment of the back-end project is completed, deploy the front-end project with the same server as above, use the "Npm run build" command to package the project, upload the package file dist to the server through Filezilla, and use NGINX as the reverse proxy to deploy the front-end project.

#### 4. System test.

**4.1. Test preparation.** In order to verify the effectiveness of the digital agricultural greenhouse remote monitoring system based on the Internet of Things proposed in this paper, the system test is conducted as shown below. First, according to the above system hardware and software design content, build a remote monitoring system. Secondly, establish the required environment for this system test. ZigBee wireless sensor network is constructed with a tree structure, which is composed of a coordinator node, a routing node, and two terminals. The terminal node is responsible for collecting environmental data in the greenhouse and sending it to the router node through the network. The routing node is responsible for gathering the collected data and transmitting it to the coordinator node, which has the responsibility of constructing the entire network and forwarding environmental data to the upper levels of computer architecture through the Internet. The 1852 square meter standard multi-span vegetable greenhouse of a vegetable base was selected for the Zigbee network, greenhouse automatic control, and remote monitoring test. The lower computer and coordinator module were installed in the greenhouse strong current control box or independent weak current control box, and connected to the back-end server through networking.

**4.2. Test results.** In order to make the system test results more visual and clear, the remote monitoring system of a digital agricultural greenhouse based on the Internet of Things proposed in this paper is set as the experimental group, and the traditional monitoring system is set as the control group. The remote monitoring results of the two systems are compared. Divide the vegetable greenhouse into six monitoring areas with the same shape and size, labeled 01 ~ 06, set temperature and humidity sensors in the divided areas, collect the temperature change data of each area in the greenhouse in real-time, and upload it to the system center. Using MATLAB simulation analysis software, take 24 hours as the monitoring cycle, count the average temperature values of each monitoring area obtained by the two remote monitoring systems, compare the temperature remote monitoring results with the actual temperature values, and determine the accuracy of the system monitoring results. The results are shown in Figure 4.1.

It can be seen from the comparison results in Figure 4.1 that the two digital agricultural greenhouse remote monitoring systems have different operating effects. Among them, after the application of the remote monitoring system of a digital agricultural greenhouse based on the Internet of Things proposed in this paper, it can be seen that the temperature remote monitoring value in the greenhouse is closer to the actual temperature value, and the monitoring deviation is small, which indicates that the remote monitoring result of the proposed method has high accuracy, and the remote monitoring effect has significant advantages.

**5. Conclusion.** In order to improve the operation function of the traditional agricultural greenhouse remote monitoring system is not perfect, the monitoring results have low accuracy, and there is a large deviation from the actual operation in the greenhouse. The present paper presents the Internet of Things (IoT) technology and proposes a design for a remote monitoring system for digital agricultural greenhouses, that is based on IoT. Through the research in this paper, the goal of remote monitoring of agricultural greenhouses has been well achieved. The remote monitoring results are close to the actual situation in the greenhouses and can obtain

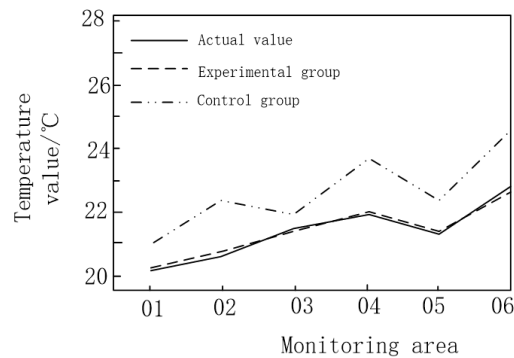


Fig. 4.1: Comparison Results of Remote Monitoring Values of Agricultural Greenhouse Temperature

the monitoring values with high accuracy, so as to grasp the dynamic changes in the agricultural greenhouses in a real-time and remote manner.

**Acknowledgments.** The 2022 school-level scientific research project of Hainan Vocational University of Science and Technology, Project No. HKKY2022ZD-09

#### REFERENCES

- [1] H. ABDOLLAHNEJAD, D. R. OCHBELAGH, AND M. AZADI, *An investigation on the 133xe global network coverage for the international monitoring system of the comprehensive nuclear-test-ban treaty*, Journal of Environmental Radioactivity, 237 (2021), p. 106701.
- [2] A. BAZGAOU, H. FATNASSI, R. BOUHARROUD, K. EZZAERI, L. GOURDO, A. WIFAYA, H. DEMRATI, F. ELAME, Á. CARREÑO-ORTEGA, A. BEKKAOUI, ET AL., *Effect of active solar heating system on microclimate, development, yield and fruit quality in greenhouse tomato production*, Renewable Energy, 165 (2021), pp. 237–250.
- [3] V. BURG, F. GOLZAR, G. BOWMAN, S. HELLWEG, AND R. ROSHANDEL, *Symbiosis opportunities between food and energy system: The potential of manure-based biogas as heating source for greenhouse production*, Journal of Industrial Ecology, 25 (2021), pp. 648–662.
- [4] Y.-C. CHEN AND H.-M. LIU, *Evaluation of greenhouse gas emissions and the feed-in tariff system of waste-to-energy facilities using a system dynamics model*, Science of the Total Environment, 792 (2021), p. 148445.
- [5] P. DONG, X. WANG, S. WANG, Y. WANG, Z. NING, AND M. S. OBAIDAT, *Internet of uavs based remote health monitoring: An online ehealth system*, IEEE Wireless Communications, 28 (2021), pp. 15–21.
- [6] N. GOEL, S. KAUR, AND Y. KUMAR, *Machine learning-based remote monitoring and predictive analytics system for crop and livestock*, in AI, Edge and IoT-based Smart Agriculture, Elsevier, 2022, pp. 395–407.
- [7] M. A. GOODWIN, D. L. CHESTER, R. BRITTON, A. V. DAVIES, AND J. BORDER, *Analysis of radionuclide detection events on the international monitoring system*, Journal of Environmental Radioactivity, 242 (2022), p. 106789.
- [8] W. T. HONEYCUTT, T. KIM, M. T. LEY, AND N. F. MATERER, *Sensor array for wireless remote monitoring of carbon dioxide and methane near carbon sequestration and oil recovery sites*, RSC advances, 11 (2021), pp. 6972–6984.
- [9] G. HU, Z. YI, L. LU, Y. HUANG, Y. ZHAI, J. LIU, AND B. YANG, *Self-powered 5g nb-iot system for remote monitoring applications*, Nano Energy, 87 (2021), p. 106140.
- [10] R. ISAENKOV, R. PEVZNER, S. GLUBOKOVSKIKH, S. YAVUZ, A. YURIKOV, K. TERTYSHNIKOV, B. GUREVICH, J. CORREA, T. WOOD, B. FREIFELD, ET AL., *An automated system for continuous monitoring of co2 geosequestration using multi-well offset vsp with permanent seismic sources and receivers: Stage 3 of the co2crc otway project*, International Journal of Greenhouse Gas Control, 108 (2021), p. 103317.
- [11] A. JENAL, H. HÜGING, H. E. AHREND, A. BOLTEN, J. BONGARTZ, AND G. BARETH, *Investigating the potential of a newly developed uav-mounted vnir/swir imaging system for monitoring crop traits—a case study for winter wheat*, Remote Sensing, 13 (2021), p. 1697.
- [12] A. C. D. S. JÚNIOR, R. MUNOZ, M. D. L. Á. QUEZADA, A. V. L. NETO, M. M. HASSAN, AND V. H. C. DE ALBUQUERQUE, *Internet of water things: A remote raw water monitoring and control system*, IEEE Access, 9 (2021), pp. 35790–35800.
- [13] J. KIM, H. LEE, S. JEONG, AND S.-H. AHN, *Sound-based remote real-time multi-device operational monitoring system using a convolutional neural network (cnn)*, Journal of Manufacturing Systems, 58 (2021), pp. 431–441.
- [14] A. MELLIT, M. BENGHANEM, O. HERRAK, AND A. MESSALAOU, *Design of a novel remote monitoring system for smart greenhouses using the internet of things and deep convolutional neural networks*, Energies, 14 (2021), p. 5045.
- [15] Y. OKADA, *Remote monitoring of railway facilities using lpwa communication*, Japanese Railway Engineering, (2022), p. 62.

- [16] M. SABAN, L. D. MEDUS, S. CASANS, O. AGHZOUT, AND A. ROSADO, *Sensor node network for remote moisture measurement in timber based on bluetooth low energy and web-based monitoring system*, *Sensors*, 21 (2021), p. 491.
- [17] Y. SHI, G. HUANG, C. AN, Y. ZHOU, AND J. YIN, *Assessment of regional greenhouse gas emissions from spring wheat cropping system: A case study of saskatchewan in canada*, *Journal of Cleaner Production*, 301 (2021), p. 126917.
- [18] S. SUJI PRASAD, M. THANGATAMILAN, M. SURESH, H. PANCHAL, C. A. RAJAN, C. SAGANA, B. GUNAPRIYA, A. SHARMA, T. PANCHAL, AND K. K. SADASIVUNI, *An efficient lora-based smart agriculture management and monitoring system using wireless sensor networks*, *International Journal of Ambient Energy*, 43 (2022), pp. 5447–5450.
- [19] J. SUWATTHIKUL, R. VANIJIRATTIKHAN, U. SUPAKCHUKUL, K. SUKSOMBOON, R. NUNTAWATTANASIRICHAI, J. PHONTIP, U. LEWLOMPHAI SARL, K. TANGPIMOLRUT, AND S. SAMRANYOODEE, *Development of dam safety remote monitoring and evaluation system*, *Journal of Disaster Research*, 16 (2021), pp. 607–617.
- [20] M. WAAGE, S. SINGHROHA, S. BÜNZ, S. PLANKE, K. A. WAGHORN, AND B. BELLWALD, *Feasibility of using the p-cable high-resolution 3d seismic system in detecting and monitoring co2 leakage*, *International Journal of Greenhouse Gas Control*, 106 (2021), p. 103240.
- [21] S. YADAV, S. PANDA, G. TIWARI, I. M. AL-HELAL, AND C. HACHEM-VERMETTE, *Periodic theory of greenhouse integrated semi-transparent photovoltaic thermal (gispt) system integrated with earth air heat exchanger (eahe)*, *Renewable Energy*, 184 (2022), pp. 45–55.
- [22] S. YAN, M. A. FAZILATI, D. TOGHRAIE, M. KHALILI, AND A. KARIMIPOUR, *Energy cost and efficiency analysis of greenhouse heating system enhancement using phase change material: An experimental study*, *Renewable Energy*, 170 (2021), pp. 133–140.
- [23] L. ZENG, C. SUN, L. ZHANG, AND ET AL., *Data collection and processing framework for operation monitoring based on big data technology*, *Computer Simulation*, 38 (2021), p. 8.
- [24] C. ZHAO, B. GAO, L. WANG, W. HUANG, S. XU, AND S. CUI, *Spatial patterns of net greenhouse gas balance and intensity in chinese orchard system*, *Science of The Total Environment*, 779 (2021), p. 146250.

*Edited by:* Sathishkumar V E

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

*Received:* May 26, 2023

*Accepted:* Jul 9, 2023







## A HYBRID EDUCATION QUALITY ASSESSMENT METHOD BASED ON HEURISTIC OPTIMIZATION ALGORITHM

**Abstract.** Hybrid education is the most effective process of education, which is the mixture of two education processes that help increase the student's understanding level. In the study, it has been discussed that the intelligent learning environment influences the education style, which allows the students and the educator to understand the topic correctly. Also, from the study, a good bond between the teacher and the students increases the achievement rate in an education system. Therefore, the professional life of the students is also more developed. The study aims to consider the disadvantages of the modern or hybrid education system, and the communication gap is one of the significant disadvantages of education.

**Key words:** Hybrid education, digital education, smart learning, heuristic optimization model, educational infrastructure.

**1. Introduction.** The word hybrid means the combination of two or more than two elements. A hybrid education system or HES is a kind of education that is a combination of the old education style and the new education style. In the present day, the culture of hybrid education has become more and more popular among the people. Hybrid education is the combination of a good quality education style that helps students for better growth and development in order to their educational life. Sometimes the old education systems are suitable for the academic growth of the students and sometimes, the modern digital education systems are better. For this reason, hybrid education is one of the best methods of education [13]. Students feel comfortable in this education system as the education system is flexible to the students. The hybrid education system is better than the monotonous and traditional education system. The hybrid system of education increases the learning opportunities of the students, and it increases their interest in learning. This increases the potentiality of the students on traditional knowledge as well as modern scientific understanding [7]. This process of education helps the students to understand the complicated matter in a short time. The clarification of the subjects is essential for a student to increase knowledge, and it helps to increase the possibility of success too.

Figure 1.1 shows the hybrid education system's process and steps. It can be said from the above picture that there are 4 factors that control the hybrid education technology: industries, government, communities, and technology. The factors of the education system are smart learning, social media, and MOOC.

Users of this education system used smart technology to take the education out of the class. The learning process of the HES is mainly the challenged-based education system and the case-based learning system [24]. The propeller planning and strategies increase the quality of the education process among the students, and the learning process becomes much easier and more comfortable. The compilation of the traditional education process and the modern education process makes education and learning helpful for educators [23].

The research on the Hybrid Quality Assessment Method (HQAM) has several potential contributions that can advance the field of education and benefit both educators and students. Here's a summary of the research contributions:

1. By evaluating the role and importance of HQAM in student growth and development, the research contributes to a better understanding of how diverse assessment methods impact learning outcomes. This knowledge can help educators design more effective assessment strategies that align with educational goals and promote deeper comprehension, critical thinking, and skill acquisition among students.
2. Through an examination of the impact of the Hybrid Education Method (HOM) on HQAM, the research identifies how the combination of in-person and online elements affects assessment results. This insight can guide educators in refining the hybrid education model to leverage its strengths, mitigate its weaknesses, and create a balanced and effective learning environment.
3. The assessment of the advantages and disadvantages of HQAM for educational development provides

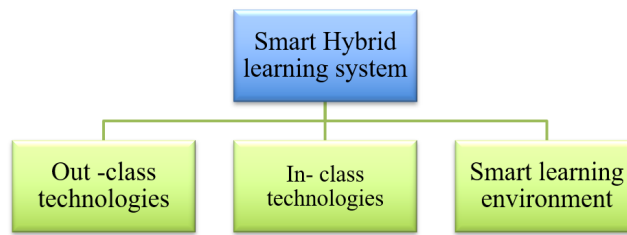


Fig. 1.1: Smart Hybrid learning system

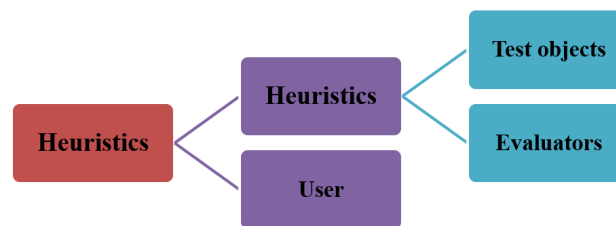


Fig. 1.2: Model of the heuristic optimization

educators with valuable information for decision-making. It empowers them to make informed choices about incorporating hybrid assessment methods in their teaching practices, considering both the benefits and challenges to maximize the effectiveness of the approach.

4. The research's exploration of the current situation of HQAM in the education process provides a snapshot of the state of implementation. This contribution highlights trends, challenges, and potential gaps in using hybrid assessment methods, facilitating discussions on how to address shortcomings and capitalize on successful approaches.

The research on HQAM makes valuable contributions to the fields of education and assessment by enhancing student learning outcomes, optimizing the hybrid education model, informing decision-making, improving assessment practices, identifying current trends and gaps, and envisioning a more effective educational future. These contributions collectively contribute to the ongoing improvement of educational experiences for both educators and students.

Heuristic optimization is the kind of algorithm that helps in the development of the students to the trial and error method. The trial and error method is one of the most effective and valuable methods for the educational growth and development of students [29]. Therefore, the proper application of the HQAM on HOM made the educational process more effective. The traditional education system was not as effective as this hybrid system of education. This helps the teachers to make the students more comfortable with different subjects. This increases the chances of getting better marks and qualifications of the students [28]. This includes the process of modern teaching process such as teaching with the help of videos and projectors to increase the imaging capability of the students. The combination of the online and traditional systems of the education process is essential and should be applicable to the development of the students [16]. This increases the probability of understanding all types of complicated things of study as well as guides the students to enhance their base of knowledge in different subjects.

The above figure shows that the model of the heuristic optimization helps in the development of the student's congenital and mental. The users are the main asset of the model and the whole model is depending on the users. Test objects and the evaluators are two factors connected with the users that help in the consideration of the model. There is a correlation between the HES and the heuristic optimization model or HOM in the case of the education system [11]. Therefore, for developing the educational infrastructure and connectivity of

the students, the HES and the HOM are two important steps.

A smart learning environment is a technologically enhanced educational setting that leverages various digital tools, data analytics, and interactive resources to create an optimized learning experience. Such environments have become increasingly important in modern education for several reasons, as they significantly influence both students and educators, ultimately demonstrating the relevance of incorporating technology in the educational process.

1. **Personalized Learning:** Smart learning environments can adapt to the individual needs and preferences of students. Through data analysis and AI-driven algorithms, these environments can identify a student's strengths, weaknesses, and learning styles. This allows educators to tailor content and activities to each student, fostering a more effective and engaging learning experience.
  2. **Engagement:** Interactive digital tools, multimedia resources, and gamified elements in smart learning environments capture students' attention and keep them engaged. This active participation helps students understand and retain complex topics more effectively than traditional passive learning methods.
  3. **Access to a Wealth of Resources:** The internet and digital platforms provide access to a vast array of educational resources, from interactive simulations to global databases of knowledge. Smart learning environments leverage this wealth of information to enhance students' understanding of subjects, allowing them to explore topics in greater depth.
  4. **Real-time Feedback:** Technology allows for instant assessment and feedback. Students can receive immediate input on their performance, helping them identify areas that need improvement and reinforcing their understanding of the material. This immediate feedback loop accelerates the learning process.
  5. **Collaborative Learning:** Smart learning environments often facilitate collaboration among students, enabling them to work together on projects, share ideas, and learn from one another. This not only enhances comprehension but also fosters essential teamwork and communication skills.
  6. **Preparation for the Digital World:** In today's technology-driven society, it's crucial for students to develop digital literacy skills. By incorporating technology in education, we prepare students for the digital challenges they'll face in their future careers.
  7. **Professional Development for Educators:** Smart learning environments empower educators by providing tools for more efficient lesson planning, data-driven insights into student performance, and innovative teaching methods. This professional development enhances educators' ability to deliver high-quality instruction.
  8. **Flexibility:** Technology enables various modes of learning, such as online courses, blended learning (combining online and in-person instruction), and flipped classrooms (where traditional homework and lectures are reversed). This flexibility accommodates different learning styles and helps students and educators adapt to changing circumstances.
- 2. Objectives.** The objectives of the research are:
1. To determine the role and importance of the hybrid quality assessment method or HQAM for the growth and development of the students.
  2. To calculate the impact of the HOM on the HQAM
  3. To estimate the advantages and the disadvantages of the HQAM for the educational development of the students in the present day.
  4. To evaluate the process of the application of the HQAM on HOM for the betterment of the students
  5. To find the current situation of the education for the application of the HQAM in the education process
  6. To elaborate on the chance of future improvement of the students applying the hybrid education system in the education culture.

**3. Proposed Methodology.** The assessment of hybrid education which is based on the HOA is one of the most impactful and effective processes that become more popular among students and teachers in the present day. The impact and both positive and negative effects of the education process have been analyzed in the study with the help of the proper methods and techniques. The issues and problems related to education techniques hamper the educational growth of educators [25]. A cognitive developmental test is inextricably linked with the heuristic optimization algorithm. All the conditions and factors the students faced were transformed into

Table 4.1: Benefits of the hybrid education process on the students

Factors of the Assesment	Benefit in the Education
Digital technology	The use of electric tools, smart classroom, and PowerPoint’s increase the cognitive development of the students.
The individual style of learning	The learning style affects the result related to the learning capacity of the students and also motivates the students.
Teacher students relationship	Encourage the students to better learning. Also helps the students to reduce depression and increase positive behavior.
Flexible education style	Flexible seating arrangements and education in the classroom made a positive effect on the students and also influenced the students in learning [31].

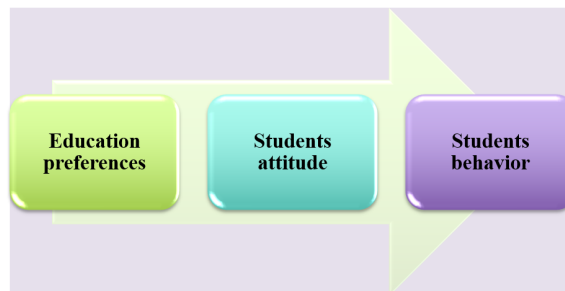


Fig. 4.1: Factors that influence the behavior of the students

sections of statements and points of vision were explored accordingly.

**4. Importance of Hybrid Quality Assesment Method.** The hybrid assessment method becomes the most valuable and profitable learning method for the better understanding of educators. The hybrid education process increases the understanding level which helps in the personal and the professional life of the students [3]. The PowerPoint presentation and the digital technology provide the students a great learning material that helps them to know more.

The use of the education process increases the quality of the education system and makes the more confident and skilled in the topic. The use of digital technology in the learning process helps the students to make clear the concept of the topic and the doubt clear season is also helpful for the students [27]. Hence, in the modern era, the HQAM made a great impact on the education process.

From the above figure, it can be said that the factors that are the reasons for the behavior of the students are the individual characteristics, needs of the students, skills, and knowledge of the students. The educational preference, the attitude of the students, and the quality of the education are also three factors that affect the behavior of the students.

The effectiveness of hybrid education in more depth:

1. **Customized Learning Paths:** Hybrid education allows for a more personalized learning experience. Students can access online resources at their own pace, review materials as needed, and explore supplementary content that aligns with their interests. This customization accommodates diverse learning styles, paces, and levels of prior knowledge, enhancing comprehension and retention.
2. **Flexibility:** One of the key advantages of hybrid education is its flexibility. Students can balance their academic pursuits with other commitments, such as work or family responsibilities. This flexibility can lead to increased motivation and reduced stress, as students have more control over when and how they engage with course materials.

3. **Active Learning:** Hybrid models often emphasize active learning strategies. With online components, students are encouraged to take a more active role in their education. This can involve interactive online discussions, collaborative projects, and hands-on activities during in-person sessions. Active learning has been shown to improve understanding and critical thinking skills.
4. **Real-world Preparation:** Hybrid education can better prepare students for the demands of the digital age. By participating in online discussions, navigating digital resources, and mastering time management skills, students develop valuable digital literacy and self-directed learning abilities that are essential in many modern workplaces.
5. **Increased Access:** Hybrid models can expand access to education. Students who may face geographical barriers, such as those in remote areas, or individuals with physical disabilities, can benefit from the online component of hybrid education, gaining access to courses they might not have otherwise been able to attend.
6. **Effective Use of Class Time:** The in-person components of hybrid education can be optimized for interactive discussions, problem-solving activities, and hands-on experiments. This maximizes the value of face-to-face time, allowing educators to focus on activities that benefit from direct interaction while shifting more traditional lectures and content consumption to online formats.
7. **Data-Driven Improvements:** Hybrid education often leverages data analytics to track student progress. Educators can use this data to identify areas where students struggle, adjust their teaching methods, and provide targeted support. This data-driven approach helps improve learning outcomes over time.
8. **Continuous Learning:** Hybrid education encourages a culture of continuous learning. Students and educators alike become more adept at using technology as a tool for ongoing education, research, and professional development.

**5. Impact of HOM on the HQAM.** The impact and the effect of the HOM on the HQAM are undeniable as the trial and error method which is used in the HOM is the most vital for the execution of the latest and modern technology in the educational institute. The modern and hybrid education system is the collection of several kinds of educational processes that help in the growth on development of students [20]. By incorporating various assessment types, the HQAM can cater to different learning styles. Some students excel in traditional written exams, while others might thrive in project-based assessments or oral presentations. This approach acknowledges the diversity of learners and promotes a more inclusive educational experience.

The importance of the heuristic optimization algorithm has been shown in the above figure where the secondary frequency control, HVDC transmission, and primary frequency control are the factors for implementing a new and latest education system. The trial and error method played a vital role. The hybrid system of learning increases the flexibility of the time period and helps the students to obtain knowledge from different sources [16]. The development of this mode of learning helps the students to get the lectures of different teachers at their homes as well and they could utilize those contents according to their own time period or schedule. This helps the students by saving their time on a journey to different situations and increasing their efficiency in knowing details of different topics [7]. In this system of education modern systems are required such as internet facilities, and digital machines such as mobile, computers or laptops.

**6. Advantages and the Disadvantages of the HQAM.** The implication of hybrid education in the educational institute attracts educators to the modern technique and technology of the education process. This education process involves the students in the hybrid education system [22]. The concept of the students becomes clearer through the service of technology in the education system.

The curriculum and teaching methodology that is used in the method is more beneficial for the students than the traditional and authentic process [4]. With the help of this education process, the relationship between the teachers and the students improves, and fear and anxiety decrease.

One of the potential limitations or drawbacks of the Hybrid Quality Assessment Method (HQAM) is the reliance on internet resources, which can pose challenges in certain situations:

**Access Disparities:** Not all students may have equal access to reliable internet resources. Socioeconomic disparities or geographical limitations could result in some students having limited or inconsistent access to high-quality online materials. This inequality can create an uneven playing field in the assessment process,

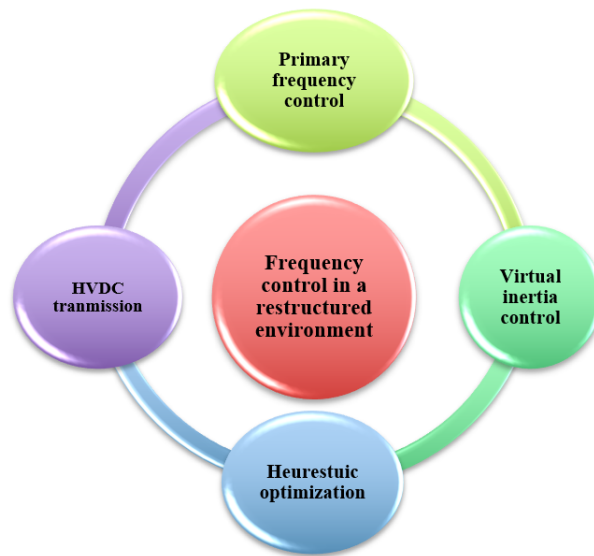


Fig. 5.1: Impact of the heuristic optimization algorithm

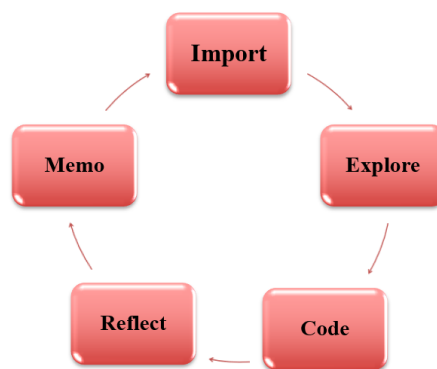


Fig. 6.1: Hybrid education model

affecting the fairness and equity of evaluations.

**Technical Issues:** Internet connectivity issues, hardware limitations, or compatibility problems with certain devices can disrupt the assessment process. Students may encounter difficulties in accessing online resources or submitting their work, leading to frustration and potential delays in the assessment timeline.

**Quality Control:** The quality and reliability of online resources can vary significantly. While the internet offers a vast array of information, not all sources are accurate, up-to-date, or credible. Depending solely on Internet resources for assessments might lead to inconsistencies in the quality of information students use, potentially impacting the accuracy of their work.

**Plagiarism Concerns:** With the abundance of online content, the risk of plagiarism increases. Students might be tempted to copy information directly from online sources without proper citation or attribution. This challenges educators to implement effective plagiarism detection mechanisms and educate students about proper research and citation practices.

**Distraction and Misdirection:** The internet can be a double-edged sword in terms of distractions. While it offers valuable resources, it also presents countless distractions, such as social media, unrelated websites, or

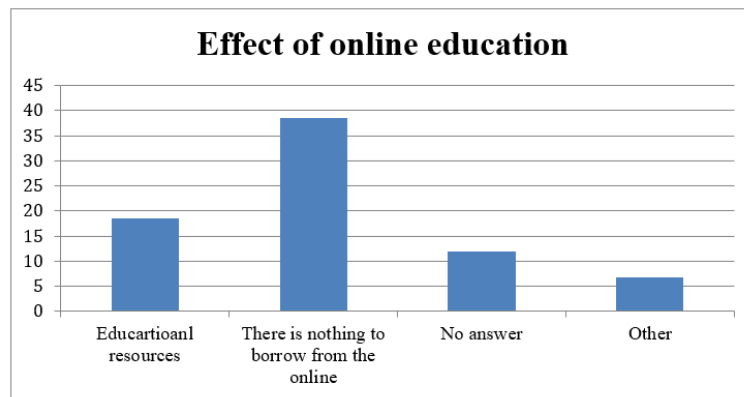


Fig. 6.2: Effect of online education

advertisements. Students might struggle to stay focused on the assessment task, impacting the quality of their work.

**Digital Literacy Requirements:** Relying on internet resources assumes a certain level of digital literacy among students. Some students, particularly those who are less familiar with technology or online research techniques, may face difficulties in navigating online resources effectively. This can hinder their ability to find and use appropriate information.

**Assessment Validity:** Depending solely on internet resources might limit the range of skills and knowledge that can be effectively assessed. Certain types of assessments, such as practical skills, hands-on experiments, or in-person interactions, may be challenging to evaluate solely through online resources.

To mitigate these limitations, educators using the HQAM should consider:

1. Ensuring that technical issues are promptly addressed and providing clear guidelines for submission.
2. Emphasizing critical evaluation of online sources and teaching proper citation techniques.
3. Encouraging focused work habits and minimizing potential distractions during assessments.
4. Offering support and resources to enhance students' digital literacy skills.

Balancing the benefits of internet resources with these potential limitations is essential to create a fair and effective assessment process within the HQAM.

Figure 6.2 has shown that the highest number of the effective field of online education is that there is no item that can be borrowed from the online which is 38.6 and the value of the educational resources is 18.5. The satisfactory and friendly relationship between the students and the teacher encourages the students to provide education quality to the teacher [8]. This becomes the best mode of education and helps the students and teachers both to increase their learning process from earlier times. On the other hand, this system is not effective for casual students because in this system all the things and learning processes depend on the learning habits of the students only [15]. Students who are not serious about their future are not adequate for this modern hybrid education system. They need to be observed by their parents and teachers and for those students traditional system will be better [12]. A hybrid education system increases the interest of students in study and it enhances the practicality of the learning process.

The use of the modern technology also leads to an increase in the ability of the trial and error method that is used in the HOM. The application of HQAM helps to select the parameters of the HOM. The use and the application of better technology also increase the capability of the machine tools and items that are used in the different kinds of sectors in the educational institute [30]. The capability of the prediction of becoming more accurate and perfect by the use of the HQAM.

The relationship between the students and the teacher becomes stronger through the use of hybrid learning in the education system as this learning process requires digital communication between the students and the teachers [6]. Around 85% of the country all over the world adopted the education system which is digital

Table 6.1: Advantages and disadvantages of the HQAM

Advantages	Disadvantages
One of the main advantages of the hybrid education strategy is that this strategy's implication of digital and modern technology helps to increase the level of understanding of the topic [19].	The limited internet sources and access is the drawback of this education system which makes the education system bounded.
With the help of this process, the students become able to think more divergently about a topic and the imagination level of the educators become increases.	Lack of interaction between the teacher and the students is the other disadvantage of the hybrid education system which makes the education process less effective.
The modern technology and the teaching process in the education system provide the educators more structural presentation of the images which helps the students to remember the topic more effortlessly [17].	The communication problem between the teacher and the students in digital learning is one of the major disadvantages [19]. Communication issues are the reason for the lack of knowledge among the students in an education institute.

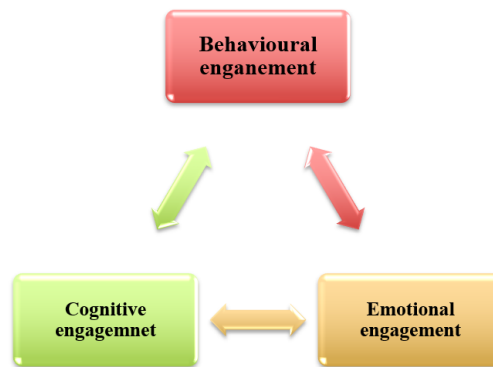


Fig. 6.3: Hybrid learning in the Student-teacher relationship

education for the effective process of the education process. The students who live far from the educational institute are becoming more beneficial for online education.

**7. Process of the Application of the HQAM on HOM.** The development of the attitude of the students is the most important and vital factor to make the students a great human being. Improving in the behavior also helps to make the development in the personal and professional life in the future [18]. The capacity for the identification of a problem and problem-solving became increased by the process. The education authority applies different types of processes and procedures to apply the HOM on the HQAM. The application of the HQAM improves the quality of education which is more effective for the cognitive and emotional development of the students.

Figure 7.2 has shown that the implication of the smart computer in the classroom enhances the quality of the education process in the modern days. The high-level strategies that are used in the HOM are effective for the growth and development of the students in their educational life [5]. The new heuristic algorithm made a great impact on the social and cognitive development of the students.



Table 7.1: Application of the HQAM on HOM

Process	Benefits
<b>Digital problem-solving process</b>	This process helps the students as well as the teachers to identify the problems and the issues systematically. The identification of the problems and issues helps to solve the problem in a better way.
<b>Smart classroom</b>	The initiative to make a smart classroom helps the teacher to describe a difficult problem in an easy way [1]. The use of the digital board can be able to show the 3d dimension picture to the students.
<b>Reviews and polling process</b>	The polling process and the review process helps to take the reviews from the students. The trial and error method or HOM also increases the accuracy of the process.
<b>Increasing the search capability</b>	The increment in the search capability also increases the boundary of the knowledge of the students.

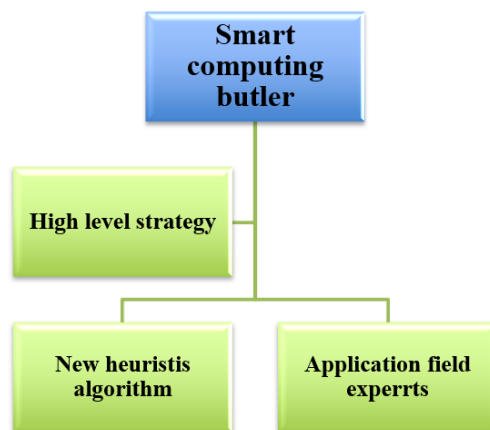


Fig. 7.1: Hybrid element – heuristic algorithm for optimizing the educational courses



Fig. 7.2: Evaluation in the computation of the HOM

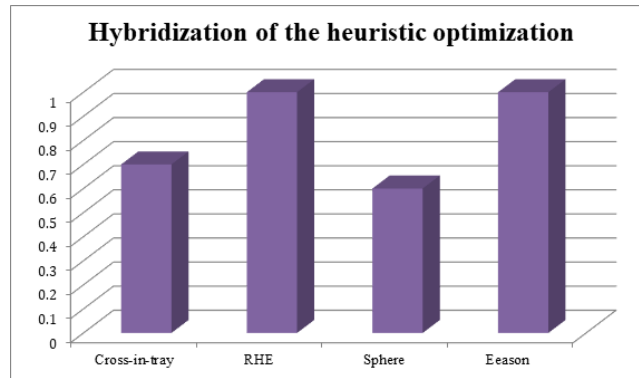


Fig. 7.3: Hybridization of the heuristic optimization in the study

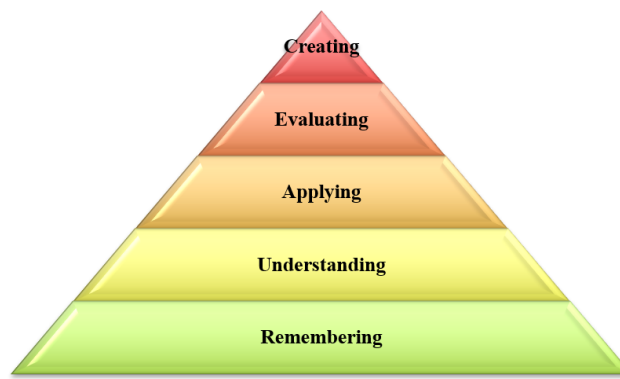


Fig. 8.1: Higher-order thinking skills

Figure 7.2 shows that the evolution of the HOM made a big impact on the educational system that increases the level of understanding of the educators. The principle of selection is one of the most essential principles that helps to select the proper issues and the problems [9]. The selection of the issues is the main key factor for reducing the problems and making the topic more clearly to the students.

The probability of the different kinds of hybridization in the education system for the various states has been shown in the above figure. From the figure, it can be said that the value of the hybridization is the highest for the RHE, MacCormik, and Eason. The value of the hybridization of the three countries is 1. The lowest rate of hybridization is for the Branin which is 0.3. The heuristic method helps to discover the attitude and behavior of the students to improve their educational careers of the students [14].

**8. Current Situation of Applying the HQAM.** Presently most of the educational institutes are focused on the application of the digital classroom and the digital education system. The use of the digital process increases the education quality and makes education more beneficial for students [21]. With the help of the HOM method, the students become able to do more practice with the help of the trial and error method [26]. The implication of this system is gradually increasing in the different education centers and schools to support the students in the learning process. The digital system of learning enhances the opportunities for learning and promotes the developmental processes of students and the whole education system [7]. The adaptation requirements of the modern education system are computerized and digitalized classrooms with internet facilities; enabling the students to use digital sources of information and helping them practice different study materials.

Table 8.1: Issues of HQAM

Issues of HQAM	Solution process
Lack of responsibilities	Among the teacher in the online education system is the reason for the lack of attraction of the student to the topic.
Scarcity of Confidence	Lack of experience and confidence among the leaders decrease the level of encouragement of the students.
Communication issues	The communication issues are the most effective for the system.

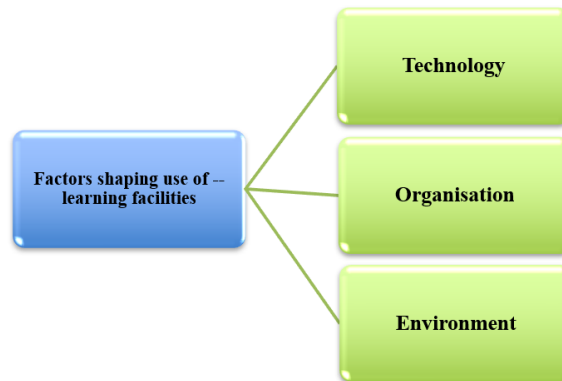


Fig. 8.2: Factors of the e-learning

The stages of higher-order thinking are shown in Figure 8.2 where the 1st stage is the stage for remembering. Understanding the topic and the proper application of the process are two vital steps according to the model. Analyzing and evaluating the strategy helps the students to make a clear view of the topic. Bloom's taxonomy is a helpful strategy to increase the cognitive development of students [10]. Therefore, it can be said that the effect and the impact of the HQAM cannot be denied on the HOM for the growth and development of the students.

The last and the final stage of the method is the creation.

Technology, organization, and environment are the three factors of e-learning which can be concluded from Figure 8.2. The e-learning facilities are mainly dependent on the educational organization, technological development and environment of the education center [12]. These need to be cooperatively developed for the development of a modernized education system.

**9. Problem Statement.** The secondary process of data collection has been used in the study to collect data about the impact of the HQAM on HOM. The numeric value of the data could be increased by the use of the primary method [2]. The numbers of resources that are used in the study are less and the use of the high number of resource can be able to make the study more effective and reliable.

**10. Results.** The application of the hybrid education quality assessment helps to increase the chance of the success of the educational industries that are based on the heuristic assessment. By the help of The hybrid quality education system or HQAM helps to increase the quality of digital learning which includes social media, smart technological learning and MOOC. The anthology of the established education process and the current education process makes instruction and learning helpful for educators. There are different ways of the importance and the essentialities of the digital and the hybrid education system which includes the enhancement in the quality of individual learning and also the relationship or the connection between the teacher and the students [4].

The above figure shows the number of students engaged online and the digital platforms in the present day.

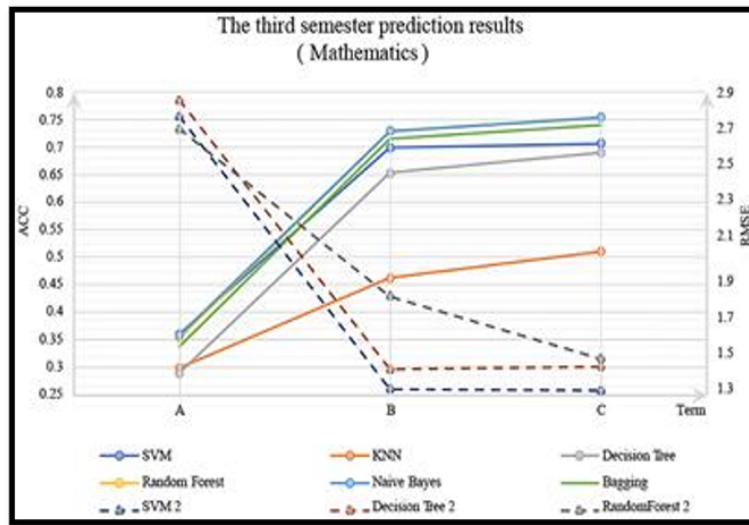


Fig. 10.1: Student engagement in the digital platform

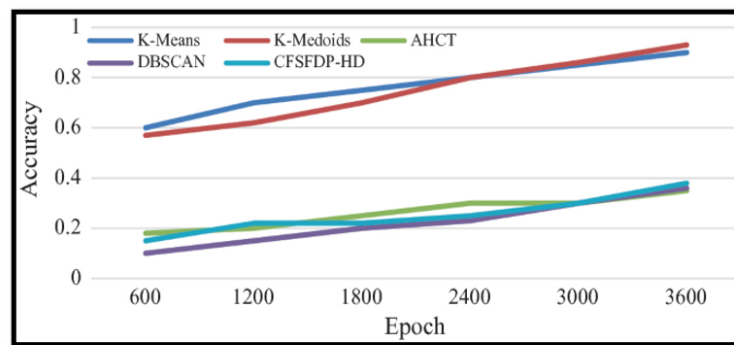


Fig. 10.2: Improving of the quality of the education system

It can be said from the picture that the number of SVM is the highest in order to the number of the prediction of the third semester. The lowest number of the students is the decision tree which is around 0.25. The impact of the HOM on the HQAM is also determined by the help of the study. The methods that include trial and error help to execute the innovative and divergent thinking of the students. The HVDC transmission that is used in the HQAM process in the mode of the frequency control of the education system which implements the latest and the modern education system. The use of modern and innovative technology is the most effective and helpful process for the educational growth and the development of the academic career of the students [13]. From the research, it can be said that the use of the HQAM enhances the knowledge and understanding of the educators.

From the figure 10.2, it can be said that the rate of the improvement of the education quality by using digital learning is the highest for the application of the K-Medoids. The value of K-Medoids is .10. The main drawback of the use of HQAM in the HOM is that the scarcity of the connection and communication between the students and the teacher through online system. Also, sometimes the rate of the understanding of the topic become impenetrable for the educators which is becomes issues for the other students. In the online class, the question answer process is also very problematic and the lack of the proper network connection lags the

process of the education.

**11. Conclusion.** In conclusion, it can be said that the application of the HQAM on HOM helps to decrease the issues of problem-solving in the classroom. The limitation in the internet resources is one of the drawbacks of the process. Different stages that are used in the HQAM process help to increase the cognitive of the students. Also, from the above study, it can be concluded that the flexible study style helps to make the students more comfortable and energetic towards the education system.

#### REFERENCES

- [1] R. ABU KHURMAA, I. ALJARAH, AND A. SHARIEH, *An intelligent feature selection approach based on moth flame optimization for medical diagnosis*, Neural Computing and Applications, 33 (2021), pp. 7165–7204.
- [2] L. ABUALIGAH AND A. DIABAT, *A comprehensive survey of the grasshopper optimization algorithm: results, variants, and applications*, Neural Computing and Applications, 32 (2020), pp. 15533–15556.
- [3] L. ABUALIGAH AND A. J. DULAIMI, *A novel feature selection method for data mining tasks using hybrid sine cosine algorithm and genetic algorithm*, Cluster Computing, 24 (2021), pp. 2161–2176.
- [4] L. ABUALIGAH, A. H. GANDOMI, M. A. ELAZIZ, H. A. HAMAD, M. OMARI, M. ALSHINWAN, AND A. M. KHASAWNEH, *Advances in meta-heuristic optimization algorithms in big data text clustering*, Electronics, 10 (2021), p. 101.
- [5] M. ASKARI AND F. KEYNIA, *Mid-term electricity load forecasting by a new composite method based on optimal learning mlp algorithm*, IET Generation, Transmission & Distribution, 14 (2020), pp. 845–852.
- [6] M. AZADI MOGHADDAM AND F. KOLAHAN, *Modeling and optimization of the electrical discharge machining process based on a combined artificial neural network and particle swarm optimization algorithm*, Scientia Iranica, 27 (2020), pp. 1206–1217.
- [7] V. BIJALWAN, V. B. SEMWAL, G. SINGH, AND T. K. MANDAL, *Hdl-psr: Modelling spatio-temporal features using hybrid deep learning approach for post-stroke rehabilitation*, Neural Processing Letters, 55 (2023), pp. 279–298.
- [8] R. DABBAGH AND S. YOUSEFI, *A hybrid decision-making approach based on fcm and moora for occupational health and safety risk analysis*, Journal of safety research, 71 (2019), pp. 111–123.
- [9] H. DING AND X. GU, *Hybrid of human learning optimization algorithm and particle swarm optimization algorithm with scheduling strategies for the flexible job-shop scheduling problem*, Neurocomputing, 414 (2020), pp. 313–332.
- [10] Y. DONG, X. MA, AND T. FU, *Electrical load forecasting: A deep learning approach based on k-nearest neighbors*, Applied Soft Computing, 99 (2021), p. 106900.
- [11] L. FANG AND X. LIANG, *A novel method based on nonlinear binary grasshopper whale optimization algorithm for feature selection*, Journal of Bionic Engineering, 20 (2023), pp. 237–252.
- [12] S. GARG, K. KAUR, N. KUMAR, G. KADDOUM, A. Y. ZOMAYA, AND R. RANJAN, *A hybrid deep learning-based model for anomaly detection in cloud datacenter networks*, IEEE Transactions on Network and Service Management, 16 (2019), pp. 924–935.
- [13] A. GOLI, H. KHADEMI-ZARE, R. TAVAKKOLI-MOGHADDAM, A. SADEGHIEH, M. SASANIAN, AND R. MALEKALIPOUR KORDESTANIZADEH, *An integrated approach based on artificial intelligence and novel meta-heuristic algorithms to predict demand for dairy products: a case study*, Network: computation in neural systems, 32 (2021), pp. 1–35.
- [14] J. HU, J. TANG, AND Y. LIN, *A novel wind power probabilistic forecasting approach based on joint quantile regression and multi-objective optimization*, Renewable Energy, 149 (2020), pp. 141–164.
- [15] B. T. HUNG, V. B. SEMWAL, N. GAUD, AND V. BIJALWAN, *Hybrid deep learning approach for aspect detection on reviews*, in Proceedings of Integrated Intelligence Enable Networks and Computing: IIENC 2020, Springer, 2021, pp. 991–999.
- [16] C. S. IRONSI AND H. B. BOSTANCI, *Utilizing mobile-learning and cap (e) lesson framework in improving the productive skills of learners in a hybrid environment*, Education+ Training, 65 (2022), pp. 232–252.
- [17] M. KHARRICH, S. KAMEL, M. ABDEEN, O. H. MOHAMMED, M. AKHERRAZ, T. KHURSHAD, AND S.-B. RHEE, *Developed approach based on equilibrium optimizer for optimal design of hybrid pv/wind/diesel/battery microgrid in dakhla, morocco*, IEEE Access, 9 (2021), pp. 13655–13670.
- [18] D. LI, S. LI, S. ZHANG, J. SUN, L. WANG, AND K. WANG, *Aging state prediction for supercapacitors based on heuristic kalman filter optimization extreme learning machine*, Energy, 250 (2022), p. 123773.
- [19] X. LIU, D. ZHANG, T. ZHANG, J. ZHANG, AND J. WANG, *A new path plan method based on hybrid algorithm of reinforcement learning and particle swarm optimization*, Engineering Computations, 39 (2022), pp. 993–1019.
- [20] X.-H. LIU, D. ZHANG, J. ZHANG, T. ZHANG, AND H. ZHU, *A path planning method based on the particle swarm optimization trained fuzzy neural network algorithm*, Cluster Computing, (2021), pp. 1–15.
- [21] A. MAOUDJ AND A. HENTOUT, *Optimal path planning approach based on q-learning algorithm for mobile robots*, Applied Soft Computing, 97 (2020), p. 106796.
- [22] B. MOHAMMADI AND S. MEHDIZADEH, *Modeling daily reference evapotranspiration via a novel approach based on support vector regression coupled with whale optimization algorithm*, Agricultural Water Management, 237 (2020), p. 106145.
- [23] F. MOSLEHI AND A. HAERI, *A novel hybrid wrapper-filter approach based on genetic algorithm, particle swarm optimization for feature subset selection*, Journal of Ambient Intelligence and Humanized Computing, 11 (2020), pp. 1105–1127.
- [24] Z. SADEGHIAN, E. AKBARI, AND H. NEMATZADEH, *A hybrid feature selection method based on information theory and binary butterfly optimization algorithm*, Engineering Applications of Artificial Intelligence, 97 (2021), p. 104079.
- [25] L. SUN, X. KONG, J. XU, Z. XUE, R. ZHAI, AND S. ZHANG, *A hybrid gene selection method based on relief and ant colony optimization algorithm for tumor classification*, Scientific reports, 9 (2019), p. 8978.

- [26] S. WAN AND Z. NIU, *A hybrid e-learning recommendation approach based on learners' influence propagation*, IEEE Transactions on Knowledge and Data Engineering, 32 (2019), pp. 827–840.
- [27] J. WANG, J. CAO, S. YUAN, AND M. CHENG, *Short-term forecasting of natural gas prices by using a novel hybrid method based on a combination of the ceemdan-se-and the pso-als-optimized gru network*, Energy, 233 (2021), p. 121082.
- [28] J. S. WEKESA, Y. LUAN, M. CHEN, AND J. MENG, *A hybrid prediction method for plant lncrna-protein interaction*, Cells, 8 (2019), p. 521.
- [29] B. YIN, C. WANG, AND F. ABZA, *New brain tumor classification method based on an improved version of whale optimization algorithm*, Biomedical Signal Processing and Control, 56 (2020), p. 101728.
- [30] X. YU, C. LI, W.-X. ZHAO, AND H. CHEN, *A novel case adaptation method based on differential evolution algorithm for disaster emergency*, Applied Soft Computing, 92 (2020), p. 106306.
- [31] S.-P. ZHU, B. KESHTEGAR, M. E. A. B. SEGHER, E. ZIO, AND O. TAYLAN, *Hybrid and enhanced pso: Novel first order reliability method-based hybrid intelligent approaches*, Computer Methods in Applied Mechanics and Engineering, 393 (2022), p. 114730.

*Edited by:* Sathishkumar V E

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

*Received:* Jun 9, 2023

*Accepted:* Aug 22, 2023



## USING MACHINE LEARNING ALGORITHMS TO DESIGN PERSONALIZED EXERCISE PROGRAMS FOR HEALTH AND WELLNESS

**Abstract.** The research paper showcases an elaborate study of machine learning, which is used in healthcare or medical platforms and can be used by healthcare professionals to adopt better diagnostic instruments and tools for examining medical issues or images. The paper highlights that a machine learning algorithm can be utilized in X-rays or MRI scans to examine disease and health issues. This paper will also discuss how this algorithm can help healthcare professionals, doctors, and nurses make accurate diagnoses for better services and patient outcomes. One of the major advantages of using the secondary research method in the following research is the abundance of the literature. All the data being used here are previously collected and evaluated with the result, and using these will increase the impact of the study overall. This method saves resources, including money, time, and manpower. This research method allows the researcher to build new knowledge and draw new conclusions based on existing expertise and knowledge. The chosen research philosophy is the Interpretivism research philosophy. The chosen research approach here is the inductive research approach. The chosen research design for this study is exploratory. All these help the research to achieve its objectives and reach the proposed goal of this research.

**Key words:** Machine learning, Drug discovery, Clinical trials, Electronic Health Records (EHRs)

**1. Introduction.** In healthcare, machine learning is a progressing field of research and development in precision medicines with several potential applications [1]. It has been noted that, as patient data is easily accessible, machine-learning algorithms in medical platforms will become highly important for healthcare experts and the overall medical system to extract accurate issues from medical information. The figure 1.1 given below shows the pillars of machine learning for the health care sector

Machine Learning (ML) has emerged as a transformative force in healthcare, revolutionizing how medical data is analyzed, interpreted, and utilized. In the realm of predictive analysis, ML plays a pivotal role in harnessing the power of data to make accurate forecasts and informed decisions. Its ability to identify various health issues and diseases provides valuable insights into the potential applications of ML in healthcare.

Predictive analysis involves using historical data to predict future outcomes or trends. This can translate into early disease detection, personalized treatment plans, and optimized patient care in healthcare. ML algorithms excel in this domain because they can recognize intricate patterns, even within large datasets, that may elude human observation. By analyzing complex relationships between variables, ML models can anticipate health risks, disease progression, and patient outcomes. ML's role in predictive analysis in healthcare is underscored by its ability to process diverse data sources. This includes electronic health records, medical imaging, genomics, wearable devices, and patient-generated data. Integrating and analyzing these data types can enable accurate predictions regarding disease susceptibility, treatment response, and potential complications. For example, ML algorithms can identify subtle changes in medical images, such as identifying early signs of cancer or anomalies in brain scans, improving diagnostic accuracy.

Furthermore, ML's predictive capabilities are instrumental in identifying health issues before they manifest clinically. By detecting subtle deviations from standard patterns, algorithms can forecast the onset of conditions like diabetes, heart disease, and mental health disorders. Early identification facilitates proactive interventions, enabling healthcare providers to initiate preventive measures and lifestyle modifications to mitigate risks. The application of ML in predictive analysis also extends to drug development and treatment optimization. ML models can predict drug efficacy and potential side effects based on patient characteristics, genetic profiles, and disease parameters. This empowers clinicians to tailor treatments, minimizing adverse reactions and maximizing

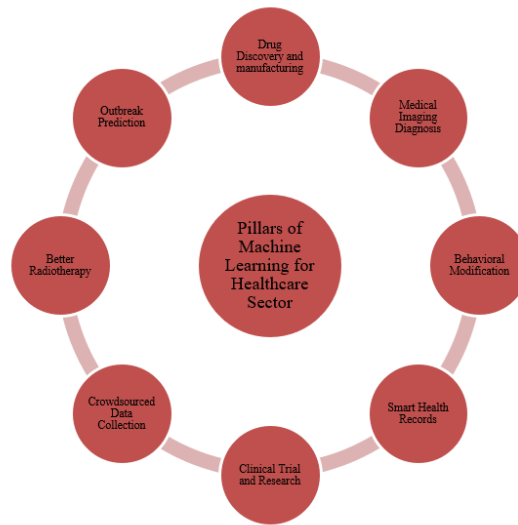


Fig. 1.1: Pillars of ML

therapeutic outcomes.

Moreover, ML's ability to process and analyze vast amounts of data leads to the generation of predictive models that continually evolve and improve over time. As new data is collected, these models refine their predictions, ensuring higher accuracy in healthcare decision-making.

In essence, this research significantly contributes to the advancement of healthcare by highlighting the transformative potential of machine learning in diagnosing medical conditions through medical imaging analysis. It underscores the importance of adopting innovative technologies to enhance healthcare practices and emphasizes the role of secondary research methodologies in consolidating existing knowledge for the generation of novel insights. The chosen research philosophy, approach, and design collectively bolster the study's objectives and contribute to its successful outcome.

**2. Objectives.** To describe the topic, the following objectives were created:

1. To discuss significant factors related to machine learning algorithms in the case of healthcare and wellness.
2. To elaborate on the role of the machine and deep learning in improving patient treatment, diagnosis, and patient outcomes.
3. To analyze the opportunities in machine learning for health.
4. To identify the key challenges related to implementing machine learning in the healthcare industry.
5. To describe the pros and cons of machine learning technologies in healthcare.

**3. Methodology.** The methodology part is one of the most important aspects of a paper that determines the different steps used to meet the study's objectives [5]. The above figure shows that the study's outcome relies on the research methods considered for the study. Besides this, secondary qualitative methods provided an extensive spectrum related to the topic of study [16]. Hence, the study is presented as an appropriate source of knowledge related to the machine learning approach as a boon to the healthcare world.

**4. Machine Learning to Develop New Treatments, Drug Discovery and Clinical Trials.** In the medical system, machine learning tools and algorithms are essential and valuable as these algorithms can help the healthcare professional to make massive amounts of healthcare data, which are created daily within electronic medical data and records [3]. With the help of machine learning tools, medical experts can find exact patterns and insights into health-related data that are difficult to find manually. It has been noted that





Fig. 2.1: Major applications of ML algorithms



Fig. 4.1: ML features for medical services

machine learning algorithms in medical systems have gained widespread adoption, and healthcare experts get the opportunity to use more predictive approaches for stratified medicine [4].

The above figure shows different features of ML technology that can be used in healthcare industries. In healthcare, the most common use cases for ML, among medical experts, are automating prescriptions, medical billing, clinical decision support, and improving clinical practices within the medical system [6]. In the clinical workflow, machine learning can produce healthcare data, and it offers primary care provider clinical decision support within medical health records [7]. Along with this, ML algorithms can be used by healthcare professionals and pharmaceutical companies to identify accurate and relevant data that could help in drug discovery, new drug development, and new treatments.

The most exciting thing about drug discoveries through AI is the convergence between drug development, biology, and technology. All these will lead to more developed medications in a more rapid time [8]. Utilizing all these effects that the ecosystem of technology, including Silicon Valley, has to offer and developed so far will benefit human society. These will leave a considerable impact on the people who are suffering from various diseases. In today's scenario, without the help of AI, the process of drug development takes almost ten years [2]. On the other hand, with the help of AI, the time can be researched to nearly one-tenth, and in this meantime,

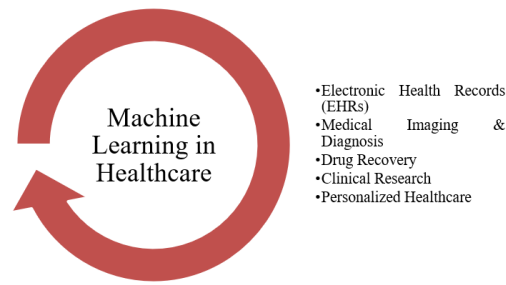


Fig. 4.2: Use of machine learning in Healthcare

the process of being found out of a drug to deliver to the patients to treat them all can be done and dusted. In case the medical research history is being studied, it can be seen that many diseases today do not have the proper treatment or drug to treat them. Fundamentally thinking, the human being can have the life-changing drugs that can change the game forever if it is delivered by the AI [20]. AI can help to reach the right patient at the right time with the newly developed drugs that no one has seen before and it can save a lot of lives.

It has been noted in the above figure that machine-learning technology in healthcare can also be used to examine the medical data from clinical trials to address earlier unknown aftereffects of drugs. However, this type of technology can help to improve patient treatment, new drug development, safety, and efficiency of medical procedures [9, 13]. Moreover, machine learning technology can also help to optimize patient treatment protocols by examining patient information, medical history, as well as real-time monitoring data.

In the future, that is not very far, it can be possible to collect all the health-related data from different types of input that are documented at different times. All these gadgets can include electronic devices to store medical data, wearable, or from research papers, whether academic or clinical [24]. Everyone will get the opportunity and access to the data that were being uploaded into secure, central storage that is trusted. It is fascinating to imagine that the AI can predict what drug to provide to what patient, and not only that it can also predict what drug will start working at exactly what time and what the sequence is [12].

According to binary classification, these algorithms can provide completely personalized care and treatment and predict patient outcomes by recognizing patient information patterns. Along with this, these algorithms allow healthcare experts and professionals to tailor the treatment process based on the individual characteristics of every patient [10, 14]. However, this led to better patient care and improved patient outcomes.

The above figure 4.3 depicts that, ML approaches offer a set of technologies that can help in new drug discovery and clinical trials with abundant and high-quality data. Clinical trials are important for new drug development and ML algorithms play a significant role in optimizing clinical trials by forecasting patient admission rates and recognizing the number of trial participants [15, 11].

The above figure 4.4 describes the clinical data of the patients. ML algorithms help to make the design of clinical trials more effective, and lower treatment costs by analyzing the previous trial data and characteristics of the patients. Besides this, machine-learning approaches enable medical experts to respond better to some specific treatments, predict the post-treatment effects, and predict the effectiveness of the treatment process [16].

However, these approaches also provide better treatment and care strategies and better patient outcomes. Thus, healthcare industries can develop innovative treatment processes, therapies, and new treatment protocols by adopting machine-learning approaches.

However, these approaches also provide better treatment and care strategies and patient outcomes. It has been noted that ML technology has the ability to transform the medical system and develop new treatment processes, drug development, and clinical trials. Thus, healthcare industries can develop innovative treatment processes, therapies, and new treatment protocols by adopting machine-learning approaches.

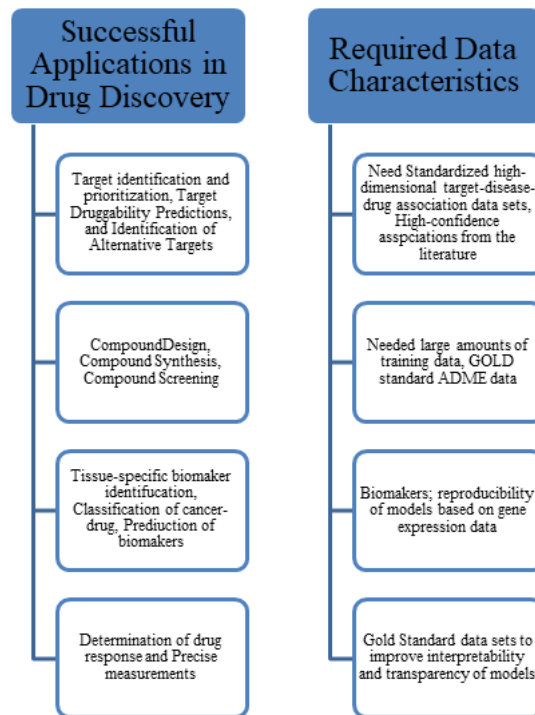


Fig. 4.3: Applications of machine learning in drug discovery and clinical traits development

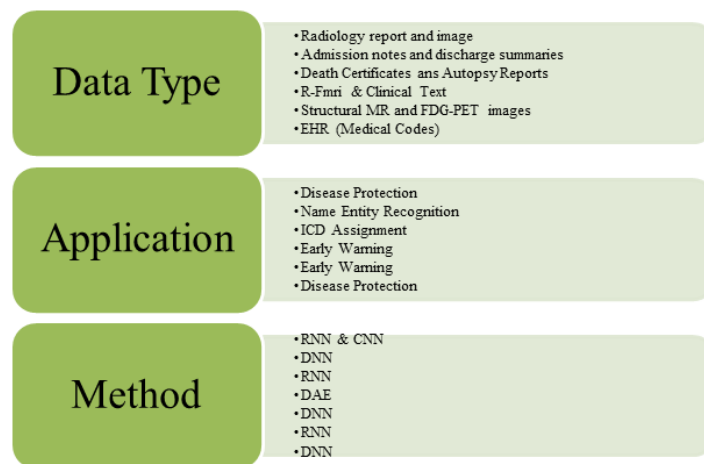


Fig. 4.4: Clinical big data and deep learning

**5. Analyze the Medical History.** In the healthcare industry machine learning, works like a blessing. In the healthcare and pharmaceutical industries, a massive amount of data is presented in the form of new and old records of patients, previous treatment as well as medical history of both patients and their families. ML approaches help to examine the previous records and data of the patients [17]. In this way, clinicians, nurses, and medical experts can predict the exact health condition of the patient.

It has been depicted in figure 5.1, that the key issues in the healthcare system such as overtreatment,

Table 4.1: Pros and cons of ML technologies in healthcare

Pros	Cons
MLO algorithms provide an accurate diagnosis and better prediction of disease outcomes.	Machine learning models are complex and there is an absence of transparency.
This allows for faster diagnosis and treatment plans for the patients.	This approach requires human oversight and surveillance.
It helps in remote monitoring of patients' medicine and improves patient convenience	Using this algorithm can create legal and regulatory issues.
Using ML technology in healthcare can provide a more effective and efficient treatment process.	ML approaches use patients' personal and sensitive data which can create ethical concerns regarding privacy and security.

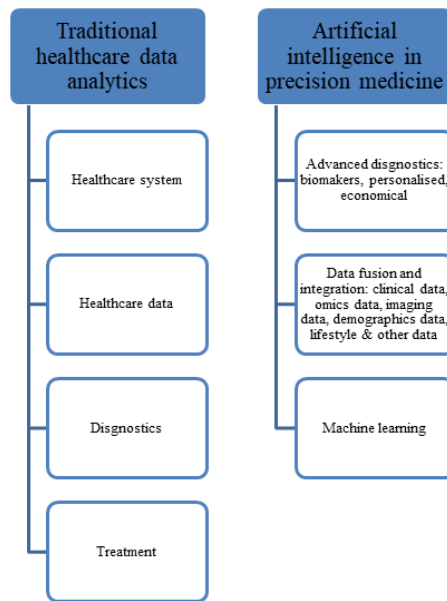


Fig. 5.1: ML in traditional healthcare data analytics

misdiagnoses, decreased productivity, and under-utilized data can be intelligently analyzed providing economic, and personalized treatment options. Along with this, ML approaches allow medical experts to extract valuable insights from previous records such as old diseases, family medical history, and hereditary diseases [18].

From the above figure 5.2, it is crystal-clear that imaging, lifestyle data fusion, logistic regression, decision tree, and discriminate analysis can be done deftly to cater to generic algorithms. ML algorithms play a significant role in pre-posing patients' records for ensuring data consistency and data privacy [19]. However, this step includes data cleaning, identifying missing data, and data transformation in a suitable format. It has been seen that ML algorithms help to extract more useful data from previous records. This process includes old medical records, patient demographics, diagnosis or test outcomes, MRI reports, chemotherapy records, and clinical notes [21]. Besides this, these technologies help healthcare experts to make more effective, efficient, and advanced treatment processes to improve patient's health.

**6. Electronic Health Records (EHRs).** EHRs comprise a surfeit of information for the patient's health history that involves both unstructured and structured data such as lab test results, disease codes, and treatments. A combination of ML and DL has been significant in maintaining the EHRs in a well-directed

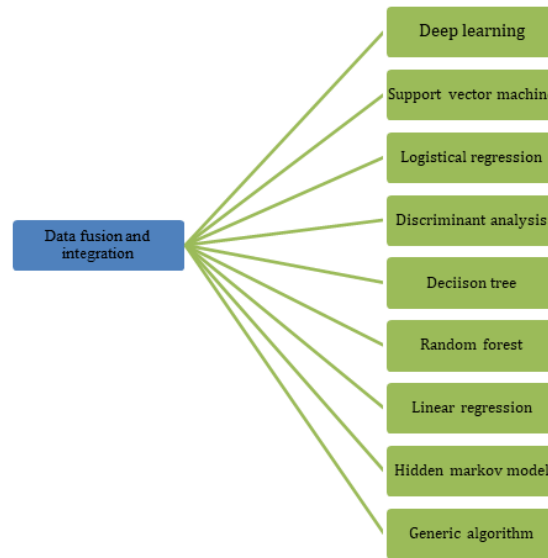


Fig. 5.2: Applying ML for genomics

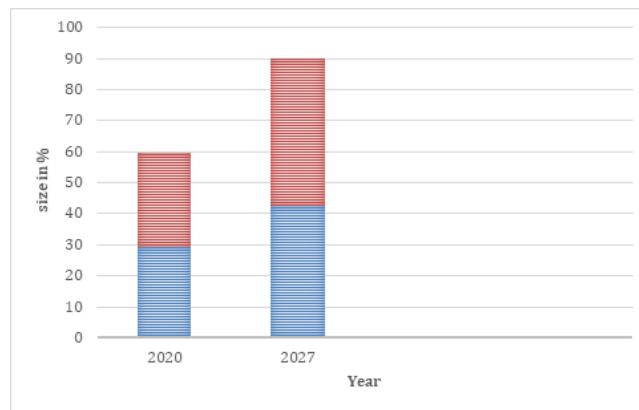


Fig. 6.1: Market size of EHR

way [22]. All these will lead to a personalized experience of getting diagnosed and getting the drug as per the requirement of the patient. In these cases, there is a slight chance of getting wrongly diagnosed but the chance is so slim [20]. Treating patients with AI can open the door to a whole new world of drugs and treatments.

From the above figure 5.2, it is writ large that the market size of EHRs is expected to reach over 47 million USD by 2027. For healthcare organizations, electronic health records are the major advancement in ML [23]. The figure reflected that the measurement was 29.16 during 2020; however, it can be accelerated within 7 years to 47.25. It has been observed that EHR allows healthcare providers to access patients’ old medical records and reports. At the same time, the same reports and records can be accessed by another doctor who wants to know which treatments and medicines have worked positively in the past.

EHRs are expected to boost the market in the near future as the market value can reach 18 billion USD by 2026 as shown in the above figure 6.2 [25]. It has been noted that electronic health records systems cover the entire array of patient records.

The overhead table 6.1 is showing the pros and cons of EHRs. Along with this, EHR is a significant

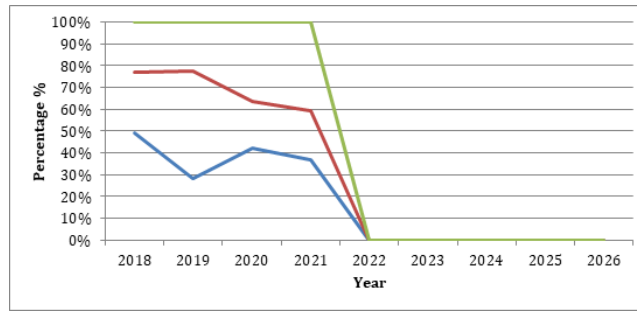


Fig. 6.2: Market value of EHRs in smart hospitals

Table 6.1: Pros and cons of EHRs in healthcare

Pros of EHRs	Cons of EHRs
EHRs provide better analysis and information.	Healthcare professionals need proper training to effectively use EHR systems.
EHR allows healthcare providers to access patient information safely and securely.	EHR may receive incorrect or incomplete information.
EHRs increase accuracy and promote transparency.	Privacy and Cyber security Issues

application of ML algorithms that can be done with the help of OCR technology. Cross-functional access and automatic document scanning can be possible through EHR [26]. Thus, smart medical records help to save time and money, in administrative record management and provide super-effective patient outcomes.

Machine learning has emerged as a transformative force in the field of healthcare, revolutionizing the way medical data is processed, analyzed, and utilized. Its importance in healthcare stems from its unparalleled ability to extract meaningful insights from vast and complex datasets, enabling healthcare professionals to make more informed decisions and deliver personalized patient care. One of the key contributions of machine learning in healthcare is its potential to enhance diagnostic accuracy. Machine learning algorithms can analyze medical images, such as X-rays, MRIs, and CT scans, with a remarkable level of precision, aiding in the early detection and diagnosis of various diseases. This not only speeds up the diagnostic process but also improves patient outcomes by enabling timely interventions.

Machine learning also plays a critical role in predictive analytics. By analyzing patient data, including medical history, genetic information, and lifestyle factors, machine-learning models can predict the likelihood of certain medical conditions or complications. This empowers healthcare providers to take proactive measures and design personalized treatment plans that align with the unique needs of each patient. Another area where machine learning shines is drug discovery and development. It can analyze massive datasets to identify potential drug candidates, predict their efficacy, and even simulate their interactions within the human body. This accelerates the drug discovery process and holds the promise of bringing new treatments to market faster.

In healthcare operations, machine learning optimizes resource allocation and improves efficiency. It can predict patient admission rates, optimize staff scheduling, and enhance inventory management, ultimately leading to cost savings and better resource utilization. Furthermore, machine learning contributes to personalized medicine by tailoring treatments to individual patients. By analyzing genetic and molecular data, machine learning models can identify the most effective treatment options and predict how patients respond to specific therapies, minimizing adverse effects and increasing treatment success rates.

In summary, the importance of machine learning in healthcare is profound. It empowers healthcare professionals with data-driven insights, improves diagnostics, accelerates drug discovery, enhances patient care, and contributes to the overall advancement of medical science. As technology continues to evolve, the integration of machine learning is expected to further transform healthcare by enabling more precise, efficient, and

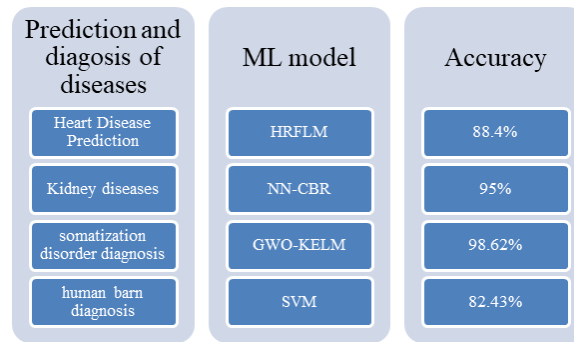


Fig. 7.1: Predict and Diagnosis of Diseases

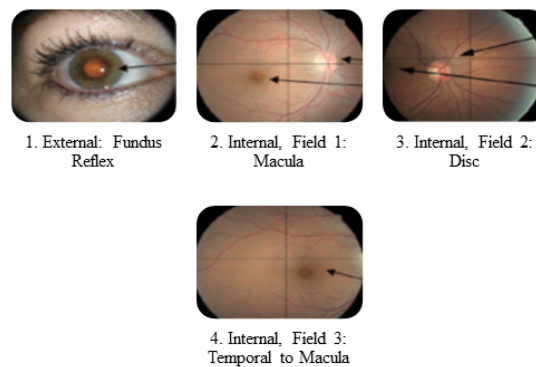


Fig. 7.2: Medical imaging for diabetic retinopathy

patient-centered approaches to diagnosis, treatment, and disease prevention

**7. Results.** In medical imaging, the aspects of ML are spearheading these days in the form of making decision trees, protein function prediction, swarm intelligence, and supporting the vector machine with evolutionary algorithms. Machine learning algorithms extract relevant information from massive unstructured data such as test reports, clinical notes, and electronic medical records [27].

The able figure depicts ongoing research in the diagnosis of different diseases. ML algorithms can monitor effective signs and symptoms including heart rate and blood pressure for identifying the early warning symptoms and signs of any health issues and diseases. ML algorithms can monitor the most vital signs by examining the video MRI reports, and scans [28].

From the overhead figure 7.2, it is clear that the screening using the *fundus photography* along with manual image analysis has been beneficial in yielding sensitivity with a heightening of the specificity rates. Medical imaging is one of the significant advancements in Machine Learning for healthcare industries [24]. Medical imaging has become more effective and advanced with the help of machine learning technology as it possesses a huge amount of pathology and radiology and makes faster decisions. It also helps doctors detect tumors at their initial stage prevent their growth and provide an accurate image of the disease to provide the doctor with the visual representation of the disease.

The overhead figure 7.3 shows that the process of setting up the CTA imaging with *computational FFR* has a significant impact on *per-vessel assessments* to maintain the quality of the diagnosis procedure.

Predictive analysis can help medical experts make informed decisions about treatment processes based on patients' health conditions. ML models help doctors to identify potential health issues, and disease management, and prevent readmissions. Predictive models can be used to identify various types of health issues and diseases

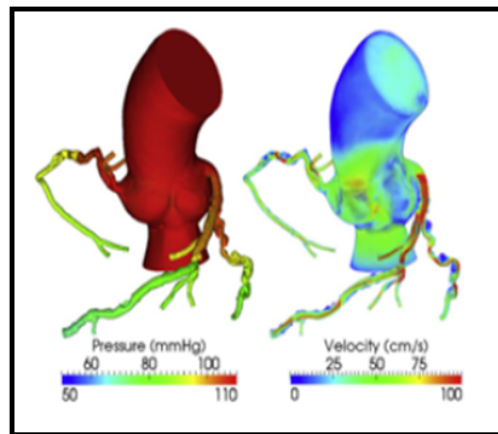


Fig. 7.3: Cardiac cycling using the CTA imaging

like heart disease, kidney disease, diabetes, liver disease, and cancer and take effective actions for patients who are in dangerous health conditions [2]. It has been noted that ML technologies can examine the data from electronic health records as well as other sources like medical records, and test results. Besides, this, by incorporating ML algorithms with predictive models, healthcare providers can predict the requirement for medical services and resources including beds, operating rooms, and staff [27].

With the help of wearable smart devices healthcare services are getting more advanced and super effective. Smart devices, tools, and instruments are resulting in a great advancement in personal health monitoring and ML technology has also taken a most important part in this. ML technology for medical systems enables completely personal as well as patient-centric treatment processes through apps [20]. There are several AI bots presented in the applications that can help to identify common problems and queries and forward the same patient to other doctors.

On the other hand, patients are able to make proper and best decisions for their health and lifestyles by uploading their health-related information in the app. Similarly, older patients can also create an emergency alarm on the smart devices because whenever they need any help, they can call for emergency help. The benefit of ML algorithms in healthcare industries for such devices is plenty. Since the patient data and information are uploaded in the online app, both patients and doctors can easily access these records to understand the real-time condition of the patient's health.

Identification of pre-signs of any health risks is another significant advantage of ML technology in healthcare industries. ML technology helps to analyze the huge amount of patient records because clinicians and healthcare providers can identify the risks before they turn to any kind of health disease. On the other hand, clinical facilities can also use this technology to identify strokes from current health conditions, heart conditions, as well as any other complications [12]. Therefore ML algorithms provide a real-time examination and investigation, and healthcare experts can diagnose potential much before.

This extreme secondary research showed that there are several aspects of machine learning-related algorithms such as impacts of different AI (Artificial Intelligence) tools, data systems on the cloud, several digital tools and services, creating smart reports, providing smart acres, and so on that can strategically be indulged within the healthcare system so that the entire healthcare can be digitized which make the delivery of treatments to the patients fast and effective. On the other hand, according to one of the objectives, the contribution related to machine and deep learning has been found to be required in today's time so that all the services to the patients such as diagnosis, treatment, and results can become more efficient. If the medical service providers can access all the required details of the patients from the smart reports system using the cloud storage system in a minute instead of asking for details from the patients which can possibly have errors or going through several notes from previous treatments physically, the entire process would be more time effective and the patient



would get better treatment. Also, the mechanical system of the concerned industry has not been updated for long; there is plenty of room for improvement in the X-ray, MRI, and other machines through the inclusion of the machine learning system that would make the results free from, manual errors.

During the entire analysis, it was noted that the adoption of ML technology in healthcare can create several problems for healthcare providers. High-quality data is highly important for testing the machine learning models. Poor data quality can negatively affect the performance of the models and fail to provide correct results. Usually, healthcare industries require more data, and making correct predictions and analyses can become complicated by the ML models when patients' records have insufficient data [8].

Similarly, while using ML algorithms in healthcare, it becomes difficult to ensure the confidentiality and security of patient data. Along with this, it was seen that there is a problem regarding the implication of ML algorithms because made by machine learning algorithms highly depend on the data [28]. It has been observed that wrong data can lead to wrong results and can harm the patient.

**8. Conclusions.** From the overall study it can be concluded that machine learning approaches technology provides a more effective framework for healthcare professionals to design personalized treatment programs for patients' health and wellness. At the same time, the paper has highlighted the benefits and disadvantages of using ML approaches in the medical system. It was found that these algorithms ensure that the treatment and diagnosis program can meet each and every patient's needs.

This research underscores the paramount importance of integrating machine learning into the realm of healthcare, particularly in the context of medical diagnostics and imaging. By harnessing the power of machine learning algorithms, healthcare professionals can significantly enhance their diagnostic capabilities, leading to more accurate and timely identification of various health conditions. This, in turn, translates to improved patient care, better treatment outcomes, and overall advancements in the healthcare industry. The findings of this study shed light on the potential of machine learning to revolutionize healthcare practices, enabling medical practitioners to make well-informed decisions based on comprehensive data analysis. The demonstrated application of machine learning algorithms in interpreting medical images like X-rays and MRI scans holds great promise for the future of medical diagnosis.

However, it is crucial to acknowledge the limitations of this research. The study primarily relies on secondary research methodologies, which might limit the depth of original insights that could have been generated through primary data collection. Additionally, while machine learning offers transformative potential, its implementation requires careful consideration of data security, privacy, and ethical concerns, which are complex and evolving issues.

For future work, further exploration is warranted to delve deeper into the integration of machine learning algorithms with other areas of healthcare, such as patient monitoring, drug discovery, and personalized treatment plans. Addressing the challenges related to data privacy, security, and ethics will also be pivotal to ensuring the responsible and effective deployment of machine learning technologies in healthcare settings. Moreover, conducting empirical research that involves real-world case studies and validation of the algorithms' performance could provide more robust insights and a clearer understanding of their real-world impact. This research serves as a stepping stone toward harnessing the full potential of machine learning in healthcare, highlighting its capabilities, limitations, and future directions. By continuously advancing our understanding and application of machine learning technologies, we can aspire to create a healthcare landscape that is characterized by precision, efficiency, and improved patient well-being.

#### REFERENCES

- [1] A. AKBARI AND R. JAFARI, *Personalizing activity recognition models through quantifying different types of uncertainty using wearable sensors*, IEEE Transactions on Biomedical Engineering, 67 (2020), pp. 2530–2541.
- [2] F. AL-TURJMAN, H. ZAHMATKESH, AND L. MOSTARDA, *Quantifying uncertainty in internet of medical things and big-data services using intelligence and deep learning*, IEEE Access, 7 (2019), pp. 115749–115759.
- [3] L. ALEXANDER III, E. MULFINGER, AND F. L. OSWALD, *Using big data and machine learning in personality measurement: Opportunities and challenges*, European Journal of Personality, 34 (2020), pp. 632–648.
- [4] A. CHATTERJEE, M. W. GERDES, AND S. G. MARTINEZ, *Identification of risk factors associated with obesity and overweight—a machine learning overview*, Sensors, 20 (2020), p. 2734.

- [5] M. CHEN AND M. DECARY, *Artificial intelligence in healthcare: An essential guide for health leaders*, in Healthcare management forum, vol. 33, SAGE Publications Sage CA: Los Angeles, CA, 2020, pp. 10–18.
- [6] A. DEPARI, P. FERRARI, A. FLAMMINI, S. RINALDI, AND E. SISINNI, *Lightweight machine learning-based approach for supervision of fitness workout*, in 2019 IEEE Sensors Applications Symposium (SAS), IEEE, 2019, pp. 1–6.
- [7] C. FEI, R. LIU, Z. LI, T. WANG, AND F. N. BAIG, *Machine and deep learning algorithms for wearable health monitoring*, in Computational intelligence in healthcare, Springer, 2021, pp. 105–160.
- [8] B. GAUR, V. K. SHUKLA, AND A. VERMA, *Strengthening people analytics through wearable iot device for real-time data collection*, in 2019 international conference on automation, computational and technology management (ICACTM), IEEE, 2019, pp. 555–560.
- [9] L. HAGEN, K. UETAKE, N. YANG, B. BOLLINGER, A. J. CHANEY, D. DZYABURA, J. ETKIN, A. GOLDFARB, L. LIU, K. SUDHIR, ET AL., *How can machine learning aid behavioral marketing research?*, Marketing Letters, 31 (2020), pp. 361–370.
- [10] M. HASSANALIERAGH, A. PAGE, T. SOYATA, G. SHARMA, M. AKTAS, G. MATEOS, B. KANTARCI, AND S. ANDREESCU, *Health monitoring and management using internet-of-things (iot) sensing with cloud-based processing: Opportunities and challenges*, in 2015 IEEE international conference on services computing, IEEE, 2015, pp. 285–292.
- [11] O. KOTELUK, A. WARTECKI, S. MAZUREK, I. KOŁODZIEJCZAK, AND A. MACKIEWICZ, *How do machines learn? artificial intelligence as a new era in medicine*, Journal of Personalized Medicine, 11 (2021), p. 32.
- [12] S. MAGHSUDI, A. LAN, J. XU, AND M. VAN DER SCHAAR, *Personalized education in the artificial intelligence era: what to expect next*, IEEE Signal Processing Magazine, 38 (2021), pp. 37–50.
- [13] A. MAHMOUDI-NEJAD, M. GUZDIAL, AND P. BOULANGER, *Arachnophobia exposure therapy using experience-driven procedural content generation via reinforcement learning (edpcgrl)*, in Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment, vol. 17, 2021, pp. 164–171.
- [14] H. MULLER, M. T. MAYRHOFER, E.-B. VAN VEEN, AND A. HOLZINGER, *The ten commandments of ethical medical ai*, Computer, 54 (2021), pp. 119–123.
- [15] J. E. MUÑOZ, S. CAO, AND J. BOGER, *Kinematically adaptive exergames: personalizing exercise therapy through closed-loop systems*, in 2019 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR), IEEE, 2019, pp. 118–1187.
- [16] E. NOSAKHARE AND R. PICARD, *Toward assessing and recommending combinations of behaviors for improving health and well-being*, ACM Transactions on Computing for Healthcare, 1 (2020), pp. 1–29.
- [17] S. ONIANI, G. MARQUES, S. BARNOVI, I. M. PIRES, AND A. K. BHOI, *Artificial intelligence for internet of things and enhanced medical systems*, Bio-inspired neurocomputing, (2021), pp. 43–59.
- [18] S. PANDEY, U. KALWA, T. KONG, B. GUO, P. C. GAUGER, D. J. PETERS, AND K.-J. YOON, *Behavioral monitoring tool for pig farmers: Ear tag sensors, machine intelligence, and technology adoption roadmap*, Animals, 11 (2021), p. 2665.
- [19] J. RAMESH, R. ABURUKBA, AND A. SAGAHYROON, *A remote healthcare monitoring framework for diabetes prediction using machine learning*, Healthcare Technology Letters, 8 (2021), pp. 45–57.
- [20] O. H. SALMAN, Z. TAHA, M. Q. ALSABAH, Y. S. HUSSEIN, A. S. MOHAMMED, AND M. AAL-NOUMAN, *A review on utilizing machine learning technology in the fields of electronic emergency triage and patient priority systems in telemedicine: Coherent taxonomy, motivations, open research challenges and recommendations for intelligent future work*, Computer Methods and Programs in Biomedicine, 209 (2021), p. 106357.
- [21] J. SAMUEL, G. M. N. ALI, M. M. RAHMAN, E. ESAWI, AND Y. SAMUEL, *Covid-19 public sentiment insights and machine learning for tweets classification*, Information, 11 (2020), p. 314.
- [22] A. SORO, G. BRUNNER, S. TANNER, AND R. WATTENHOFER, *Recognition and repetition counting for complex physical exercises with deep learning*, Sensors, 19 (2019), p. 714.
- [23] A. THIEME, D. BELGRAVE, AND G. DOHERTY, *Machine learning in mental health: A systematic review of the hci literature to support the development of effective and implementable ml systems*, ACM Transactions on Computer-Human Interaction (TOCHI), 27 (2020), pp. 1–53.
- [24] L. TONG, H. MA, Q. LIN, J. HE, AND L. PENG, *A novel deep learning bi-gru-i model for real-time human activity recognition using inertial sensors*, IEEE Sensors Journal, 22 (2022), pp. 6164–6174.
- [25] G. M. WEISS, J. L. TIMKO, C. M. GALLAGHER, K. YONEDA, AND A. J. SCHREIBER, *Smartwatch-based activity recognition: A machine learning approach*, in 2016 IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI), IEEE, 2016, pp. 426–429.
- [26] H. YU, E. B. KLERMAN, R. W. PICARD, AND A. SANO, *Personalized wellbeing prediction using behavioral, physiological and weather data*, in 2019 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI), IEEE, 2019, pp. 1–4.
- [27] S. ZHANG, Y. LI, S. ZHANG, F. SHAHABI, S. XIA, Y. DENG, AND N. ALSHURAF, *Deep learning in human activity recognition with wearable sensors: A review on advances*, Sensors, 22 (2022), p. 1476.
- [28] J. ZHU, D. H. DALLAL, R. C. GRAY, J. VILLAREALE, S. ONTAÑÓN, E. M. FORMAN, AND D. ARIGO, *Personalization paradox in behavior change apps: lessons from a social comparison-based personalized app for physical activity*, Proceedings of the ACM on Human-Computer Interaction, 5 (2021), pp. 1–21.

*Edited by:* Sathishkumar V E

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

*Received:* Jun 9, 2023

*Accepted:* Aug 22, 2023



## ANALYZING SPECTATOR EMOTIONS AND BEHAVIORS AT LIVE SPORTING EVENTS USING COMPUTER VISION AND SENTIMENT ANALYSIS TECHNIQUES

**Abstract.** It makes the reflection of humans' emotions and intentions from watching live sports events. Watching the event keeps people entertained and changes their mindset from being stressed to joyful. Watching sports events encourages the athletes and the sports persons to participate. Reflection of the live sports event consists of many crowds as the event's audience. This crowd's emotions and intentions directly impact the changes in the event's performance. It provides positive energy to the demotivated sports participants, making them perform better in the event. This study reflects the understanding of the facial emotions of the spectators from the live event. Then, they are decoded in the computer programming language, and an outcome is provided. It understands the emotions and sentiments of the people that affect the event's environment. The representation by the computer analysis makes the understanding of the changes provided by the spectators of the live event. The effect of the audience's emotions and behaviors in the crowd are computed by the utilization of computer software analysis and the effect of those reactions in the event. The collection of data is taken from the secondary sources of data collection, including the collection of information from the article and the journal based on the topic. The gathered data is analyzed by comparing them with their reaction and expressions in the live sports event.

**Key words:** Audience, Computerized analysis, Facial expression, emotions and behaviors, sentiment analysis techniques

**1. Introduction.** The work reflects humans' emotions and intentions from watching live sports events. Watching the event keeps people entertained and changes their mindset from being stressed to joyful. Watching sports events encourages the athletes and the sports persons to participate. In reflection of the live sports event it consists of many crowds as the event's audience. This crowd's emotions and intentions directly impact the changes in the performance of the event [11]. The cheering and the crowd's support work as the key encouragement to the athletes and the sportsperson participating in the event. Therefore, it is a crucial factor for the participants' performance. It provides positive energy to the demotivated sports participants, making them perform better in the event. This study reflects the understanding of the facial emotions of the spectators from the live event. Then, they are decoded in the computer programming language, and an outcome is provided. It understands the emotions and sentiments of the people that affect the event's environment. The final representation by the computer analysis makes the understanding of the changes provided by the spectators of the live event [6]. The effect of the emotions and the audience's behaviors in the crowd are computed using computer software analysis. The effects of those reactions in the event are shown in this study.

The major contributions of this work are:

1. The study introduces the concept of facial emotion recognition to understand spectators' emotions. By decoding these emotions through computer programming, the research contributes to a method for objectively measuring audience sentiment.
2. The research showcases how spectators' emotions and sentiments directly impact the sports event's overall environment. This understanding offers insights into creating a more engaging and interactive atmosphere for athletes and the audience.
3. By utilizing computer software analysis, the research demonstrates a method for quantitatively assessing the effect of audience emotions and behaviors during live sports events. This approach provides insights into the audience's reactions and their implications.
4. The research contributes to methodological approaches by utilizing secondary sources of data collection, such as articles and journals, to gather information about audience reactions. The comparison of gathered data with live event expressions enriches the analysis.

**Importance of Emotion Recognition of Spectators.** Emotion recognition of spectators, especially during events like live sports, holds significant importance due to its various applications and implications.

Understanding the emotions of the audience can provide valuable insights that impact multiple domains:

1. **Enhancing Viewer Experience:** Emotion recognition enables event organizers, broadcasters, and content creators to tailor their offerings based on the audience's emotional responses. This customization can lead to a more engaging and immersive experience for viewers.
2. **Content Personalization:** By recognizing emotions, content can be personalized to match the preferences and sentiments of the audience. For example, during live sports broadcasts, emotional cues can trigger instant replays of exciting moments, enhancing the viewing experience.
3. **Real-time Feedback:** Emotion recognition provides real-time feedback on how the audience responds to different aspects of the event. This feedback can guide event organizers in making on-the-fly adjustments to keep the audience engaged.
4. **Understanding Audience Preferences:** Analyzing emotional patterns over time helps event organizers understand which segments of the event resonate most with the audience. This information informs decisions about future content and event planning.
5. **Enhancing Athlete Performance:** Emotion recognition can impact athletes too. Understanding the crowd's emotions can motivate athletes and influence their performance. Positive reactions can boost confidence, while negative reactions can indicate areas for improvement.

**Applications of Emotion Recognition of Spectators.** The applications of emotion recognition of spectators are broad and diverse, spanning across various industries:

1. **Entertainment and Sports:** In live sports events, emotion recognition can be used to gauge audience reactions to different game situations, players' performances, and overall event dynamics. This data can shape commentary, replays, and content creation.
2. **Market Research and Advertising:** Advertisers can use emotion recognition to assess viewer reactions to commercials and advertisements. This data helps tailor marketing campaigns to resonate better with the target audience.
3. **Healthcare:** Emotion recognition has applications in mental health monitoring. It can help identify individuals experiencing stress, anxiety, or other emotions, facilitating timely interventions.
4. **Education:** In online learning environments, emotion recognition can gauge students' engagement levels and adjust instructional content accordingly. This promotes effective learning outcomes.
5. **Human-Computer Interaction:** Emotion recognition can enhance the interaction between humans and machines, making user interfaces more intuitive and responsive. For instance, virtual assistants can adapt their responses based on users' emotional states.
6. **Customer Experience:** Businesses can use emotion recognition in customer service interactions to assess customer satisfaction and tailor responses accordingly, leading to improved customer experiences.
7. **Security and Surveillance:** Emotion recognition can be used in security systems to detect suspicious behaviors or emotional states that may indicate potential threats.

**2. Objectives.** This study properly examines and elaborates on some of the basic objectives. This includes the basic concept of the computer technology used in analyzing the emotions and the intentions of the spectator from the live sports event. That provides for the understanding of the effect of audience behaviors during the watching of the live event is a very crucial factor for the performance of the participant [7]. It provides positive energy to the demotivated sports participants, making them perform better in the event. This study reflects the understanding of the facial emotions of the spectators from the live event. Then, they are decoded in the computer programming language, and an outcome is provided. It understands the emotions and sentiments of the people that affect the event's environment. Some of the objectives of using the techniques of digital computer vision and software-based sentiment analysis in the making of the sports performance development in the sports event are as follows:

1. To elaborate on the concept of the live sports event cheering by the crowd
2. To examine some of the impacts of intention and the expression of the audience from a live sports event
3. To state some of the challenges faced due to the changes in emotions and expression in the live sports event

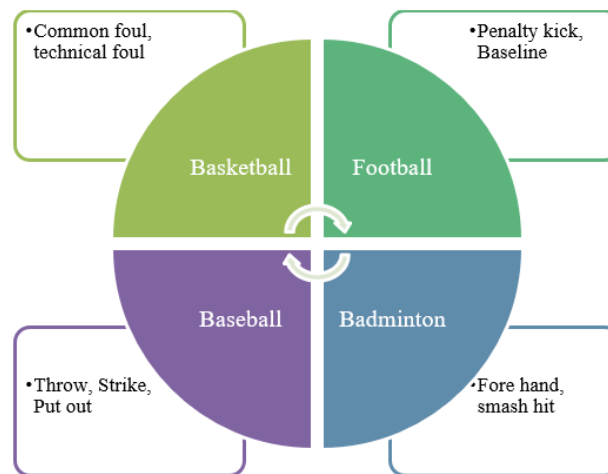


Fig. 4.1: Encouragement by the audience in the sports event

4. To implement techniques of computer-based facial expression analysis for examining the event's audience intentions
5. To describe some of the methods of computer analysis for the examination of the facial expression and intention of Human
6. To state some of the effects of computer face expression analysis on making crucial changes in the environment of the event

**3. Methodology.** In the methodology section, the study collects information about the impact of the audience in the sports event. This work represents all the collection of data from the effect of human intention in the sports event and the examination of those expressions with computer technologies for solving the problems occurring in an environment of the live event, especially in the live sports event. This work makes the presentation of all the positive effects of computer technologies. It examines the behaviors of humans during live sports events and the technological factors affecting the improvement of problems created in the event's performance and getting the visual solution [1]. The representation by the computer analysis makes the understanding of the changes provided by the spectators of the live event. The effect of the audience's emotions and behaviors in the crowd are computed by the utilization of computer software analysis and the impact of those reactions in the event. Data collection is taken from the secondary sources of data collection, which includes the collection of information from the article and the journal based on the topic [9]. The gathered data is analyzed by comparing them with their reaction and expressions in the live sports event. This includes observing the impact of computer analysis for understanding the facial expressions of the crowds in sports events.

As this study was done based on the secondary qualitative data collection method, it has helped a lot by giving various hints related to the topic. Previous articles are available beforehand, which help identify the portion that still needs more definition [10]. In various journals, scholarly articles were available there that were done by other researchers. In this way, it was easy to identify the covered portion, and after collecting the information, the data was needed to analyze, which is one of the essential parts.

**4. Encouragement by the crowd in the sports event.** The performance of the crowd changes the intention of the players and the participants of the sports event, as shown in figure 4.1. This includes the performance change of that particular sport on the field [3]. The performance of the sports persons in the event depends on the player's mental stability, concentration, and confidence during the event. The above-represented Figure 4.1 shows the impact of the cheering and the crowd's support as the key encouragement to the athletes and the sports person participating in the sports event.

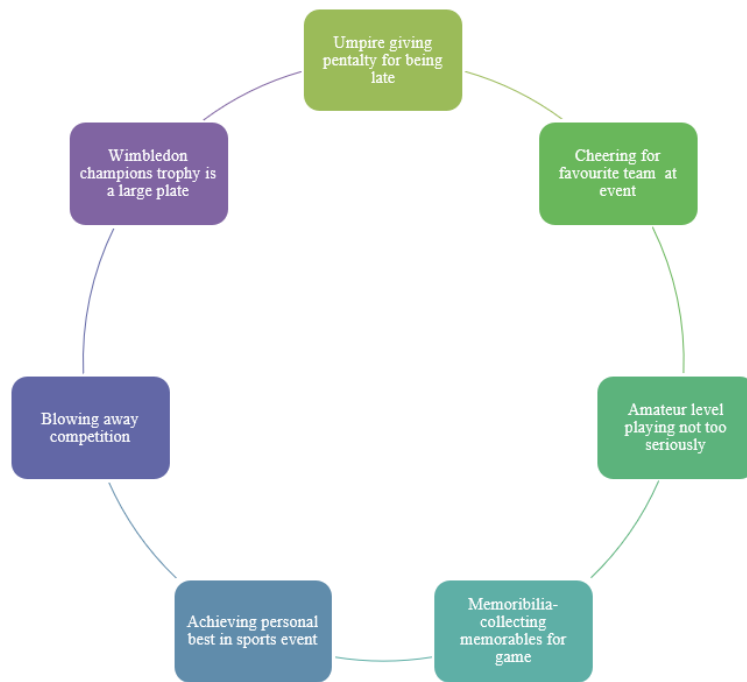


Fig. 4.2: Encouragement by the audience in the sports event

**Recognition of Input Data:** Recognition of input data refers to the process of identifying and categorizing information from various sources or inputs. This process often involves applying algorithms and techniques to analyze data and extract meaningful patterns, features, or characteristics. In the context of your abstract, recognizing facial emotions from spectators during live sports events involves capturing facial expressions, processing these visual cues, and determining the emotions conveyed by individuals in the crowd.

For example, in the study mentioned, the recognition of facial emotions could involve using computer vision techniques and machine learning algorithms to analyze video footage of the crowd. These algorithms might detect facial landmarks, expressions, and other visual cues to classify emotions such as happiness, excitement, or disappointment. This recognition process allows researchers to quantify and understand the emotional dynamics of the audience.

**Usage of Multicasting:** Multicasting is a communication technique used in computer networks to transmit data from one sender to multiple recipients simultaneously. Unlike broadcasting, where data is sent to all devices on the web, multicasting targets a specific group of recipients who have expressed interest in receiving the data. This method benefits scenarios like streaming live sports events to a large audience.

The above figure 4.2 shows that Sometimes some of the distractions and the under-confidence provided by related factors make the performance quality low. In that case, creating stability in the player's version becomes crucial. The encouragement of the crowd is the key factor in increasing the player's confidence level [16]. It provides the constructing mental stability to the participant for making a more effective decision in the event. It also creates positivity among the players and tends them to make better team collaboration for better-qualified performance.

Table 4.1 shows the necessity and the impact of encouragement by the crowd audience in making environmental changes to the live event. That includes the change of the performance of the sports persons in the event depending on the mental stability, concentration, and confidence of the player by the audience's motivation and cheering.

Table 4.1: necessity and impact of crowd encouragement in a sports event

Necessity	Impact
Making effective decisions for the event	Providing of the constructing mental stability to the participant for making a more effective decision in the event
Collaboration of team members	It also makes positivity among the players and tends them to make better team collaboration for better-qualified performance

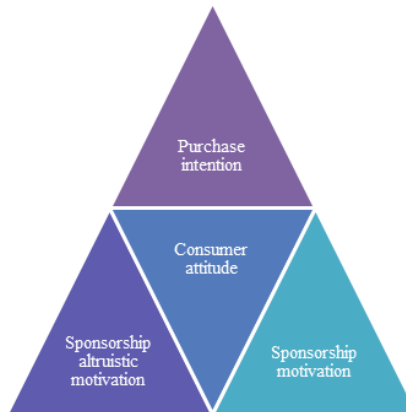


Fig. 5.1: Impact of crowd intention in making changes in the live sports event

**5. Impact of crowd intention in making changes in the live sports event.** The benefits of the crowd cheering help in making the confidence level of the participants in the sports event. For example, in the live sports event of a cricket match, the environment can be changed by the audience’s expression and intention. That sports event consists of 11 players of the team fielding on the ground and two players batting from the opponent team [13]. That becomes a mental condition of the player to be changed by the opponent fielders. In that condition, the crowd’s audience’s support and cheering are the key encouragement in providing confidence to the batsman. That makes them perform better in the match. The above figure 5.1 shows the changes in the development of the sports participant involved in the event [8]. It provides positive energy to the demotivated sports participants, making them perform better in the event. These behaviors and intentions of the crowd reflected the understanding of the facial emotions of the performer from the live event. Therefore, it can be stated that it is one of the key factors in constructing better decision-making ability for the performer of the sports event that gets reflected in their performance on the ground.

The intention of crowd to make changes in live sports events can have a positive impact from various sides. The crowd is helping to encourage the players, and along with that, it will create a positive environment [14]. The spectator’s presence always motivates and boosts the players’ morale and enhances the overall performance. As much as the spectators get interested from the match, they will visit repeatedly, enhancing the fans’ experiences. They can encourage their favorite players by waving flags, cheering, showing banners, and many more ways. The crowd can enhance the player’s mind and influence them for better performance [4]. It is important to remember that their safety and security matter most while crowds are increasing. It is one of the duties of the organizers to maintain every side.

**6. Challenges faced due to the changes in emotions and expression of the crowd in the live sports event.** Since it is known the encouragement of the crowd acts as the key factor in increasing the confidence level of the player. It makes providing of the constructing mental stability to the participant for making a more effective decision in the event. It also makes positivity among the players and tends them to make better team collaboration for better-qualified performance. It also faces many challenges that are reflected

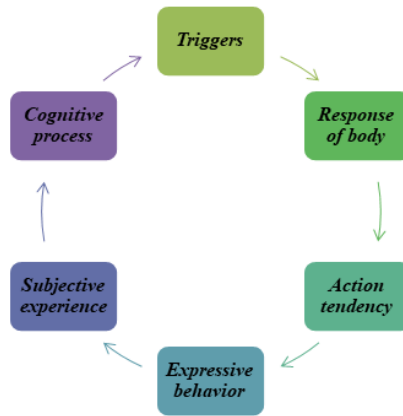


Fig. 6.1: Challenges faced due to the changes in emotions and expression of the crowd in the live sports event



Fig. 6.2: Challenges faced due to the changes in emotions and expression of the crowd in the live sports event

in the performance of the participants of the sports events [5]. As the support and the cheering from the crowd audience act as the key encouragement in providing confidence, it can also make the supporting participant get demotivated and get distracted from the way to success. The above figure 6.1 shows the challenges the sports participant faces during the live sports event. That includes the performance low-quality performance of the participants [2]. In the case of team performance in a sports event, the team members' collaboration towards getting competitive advantages is crucial. Figure 6.2 shows the better version of one player in the team cannot make the whole team perform better. The under-confidence of the team members cannot make the team achieve more competitive performance. The nervousness and the instability of the mindset act as the key challenge in demoting the better version of the team.

The above table 6.1 shows the challenges and the impacts of the challenges faced by the changes in the intention and expression of sports participants. This includes the changes in the environment of the sports event. The diverse mindset and low confidence level of the sports participants directly impact this.

There are several challenges have been faced due to the changes in emotions and expression of the crowd in the live sports event. One of the most significant risk factors is safety, as overcrowded places with minimal military support can be the reason for life risk. Various suspicious activities can happen then, and sometimes it becomes quite challenging to monitor those sides[17]. While it becomes difficult for the management to handle the crowd, it can be the reason for adverse interactions and loss of enjoyment. It is also the reason



Table 6.1: Challenges and impacts of challenges in crowd encouragement in sports events

Challenges	Impact
Mismatching of team members' collaboration	Team members cannot make the team perform more competitively performance.
The nervousness and the instability of the mindset of the sports event participant	It becomes a key challenge in demoting the better performance of the team

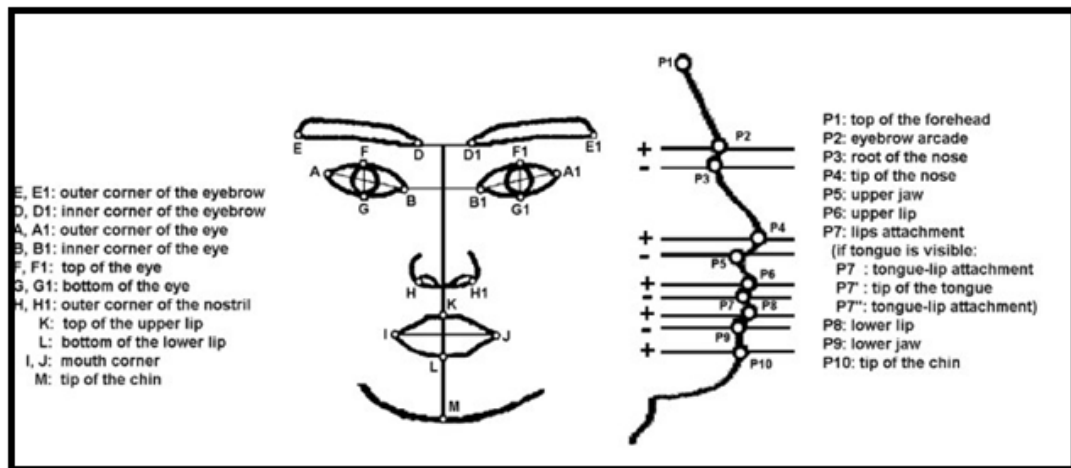


Fig. 7.1: Facial analysis of human expression by computer-based program

for reputation loss as many spectators are available in the field who intentionally spread negative information about any part.

**7. Computer-based facial expression analysis for examining the event's audience intentions.**

The crowd's encouragement provides positive energy to the demotivated sports participants, making them perform better in the event. This examination of the crowd's emotions reflects the understanding of the facial emotions of the spectators from the live event. The collection of expressions and emotions is collected digitally. Then, they are decoded in the computer programming language, and an outcome is provided [4]. It understands the emotions and sentiments of the people that affect the event's environment. The final representation by the computer analysis makes the understanding of the changes provided by the spectators of the live event [9]. The effect of the emotions and the audience's behaviors in the crowd are computed using computer software analysis. The impact of those reactions in the event is shown in this study. These all are represented in the above figure 7.1, including the behavioral changes of the audience.

The above table 7.1 shows the need and importance of computer-based analysis for examining the sentiments and emotions of the crowd during the watching of a live sports event. These expressions are collected and reflect the impact of those making the environmental changes

**8. Methods of computer analysis for the examination of the facial expression, emotion and intention of Human.**

The decoding of the collected data in the computer programming language and an outcome is provided. It makes us understand the emotions and sentiments of the people that affect the event's environment. The last representation by the computer analysis makes the understanding of the changes provided by the spectators of the live event. The computer-based analysis methods represent the digital analysis of the mindset of participants and the crowd. Examination through an automatic emotion detector is typically performed by measuring various parameters of the human body or types of impulses done through

Table 7.1: Necessity and impact of crowd encouragement examined digitally through computer software

Necessity	Impact
Examination of the crowd’s emotions reflects the understanding of the facial emotions	The final representation by the computer analysis makes the understanding of the changes provided by the spectators of the live event
Digital collection and understanding of humans emotions	It also makes the understanding of the behavioral changes of the audience.

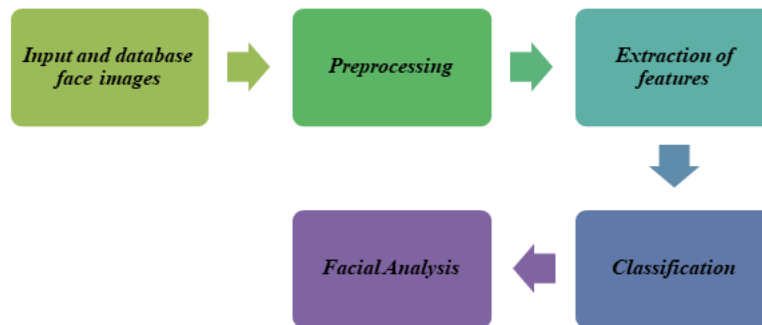


Fig. 8.1: Methods of computer analysis for the examination of the facial expression, emotion, and intention of Human

electric waves in the nervous system. It also examines the changes in intention and expression [7]. The majorly preferred methods of reviewing the choice and expression of humans are electroencephalography, skin resistance measurements, blood pressure, heart rate, eye activity, and motion analysis. The above figure 8.1 shows the methods of examining information collection in developing emotions that can hamper the health condition of the audience and the sports event participants [12].

**9. Impact of computer face expression analysis on making crucial changes in the environment of the live event.** The cheering and the empowerment of the crowd change the intention of the players and the sports event participants. This includes the performance change of that particular sport on the field. The performance of the sports persons in the event depends on the player’s mental stability, concentration, and confidence during the event. Sometimes, some of the distractions and the underconfidence provided by related factors make the quality of the performance low. In that case, creating stability in the version of the player becomes very crucial [15]. The examination of the crowd’s emotions reflects the understanding of the facial emotions of the spectators from the live event. This helps in getting the test of all the encouragement by the audience to the sportsperson participating in the sports event. The collection of expressions and emotions is collected digitally as represented in figure 9.1.

Then, they are decoded in the computer programming language, and an outcome is provided. It understands the emotions and sentiments of the people that affect the environment of the event [16]. The above figure 9.2 shows the crowd’s facial expressions and feelings are important in the communication process as they produce a positive or negative energy to the participant in the live event. It helps to understand the collector gathers significant information about the expression and the changes. In the case of collecting the face of the audiences of the live sports event it makes the understanding of the factors that can change the environment and the performance of the participants in sports events. The application of digital methods makes the representation of actual intention, and the crowd’s behavior impacts the live activity sportspersons involved in the sports event.

Table 9.1 shows the causes and impacts of analyzing the facial expression of the crowd attending the



Fig. 9.1: Impact of Computer face expression Analysis for analyzing the facial expression of human

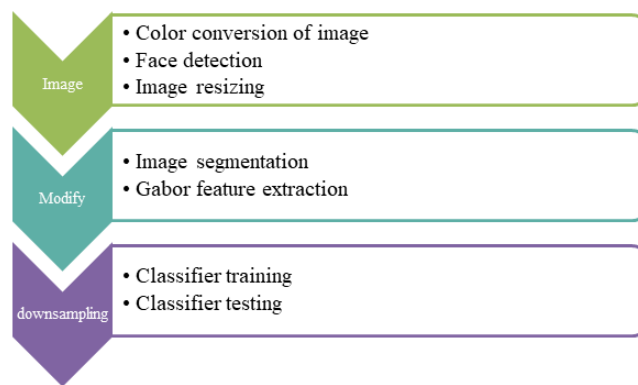


Fig. 9.2: Computerized facial expression process

live sports event. This includes the collection of information through digital human presentation examining techniques.

Automatic facial expression analysis is vital as it can identify the 'human-computer interaction' and it has become an exciting area over the past decades.

**10. Results.** Crowd emotions play a crucial and essential role in the outcome of sports games in all levels. The emotions connected to particular sports are very big to measure. Still, from research statistics it can be said that different ranges of emotions that are displayed by the different age group audiences heavily affect the sporting events.

From figure 10.1 it can be seen that between different sports fan attendances millennial age group attendance is the most. The millennial age group people are the ones who were born in the 1990s, and they are the most active sports fans throughout sporting events because they were born and brought up in an era where the internet wasn't that popular. However, Gen Z shows the second-largest attendance rate in different sporting events. Crowd emotions play a crucial and essential role in the outcome of sports games at all levels. The feelings connected to particular sports are enormous to measure. Still, from research statistics, it can be said that different ranges of emotions displayed by the other age group audiences heavily affect sporting events. The emotions shown by a higher number of audiences will also impact the players and might be instrumental in deciding the game's outcome.

From figure 10.2 it can be said that from the year 2012 to 2022, the numbers of frequent viewers, sports fans, and occasional viewers have altered quite a lot. Regular viewers are genuine sports lovers who visit the live events or watch on online platforms and make time for these events despite having a busy jobs or responsibilities.

Table 9.1: Causes and impacts of analyzing the facial expression of the crowd attending the live sports event

Causes	Impact
Collecting facial expressions and the emotions of the crowd	They help produce positive or negative energy to the participant in the live event it helps to understand the collector gathers the significant information about the expression.
collecting the expression of the audiences of the live sports event	It makes the understanding of the factors that can change the environment and the performance of the participants in sports events.

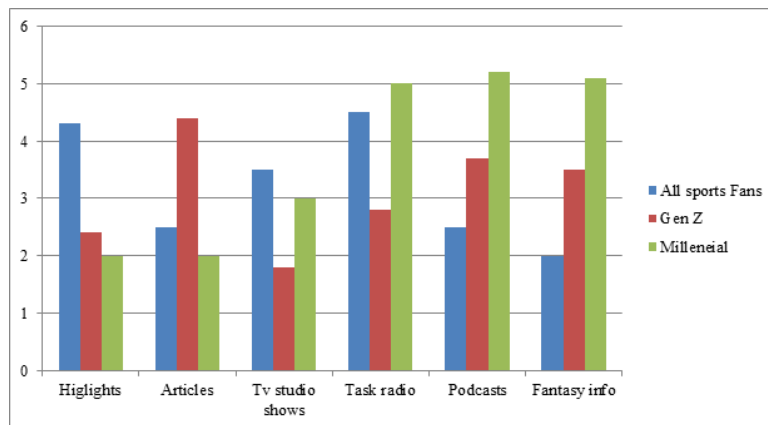


Fig. 10.1: Sports fan viewers from different generations

Occasional viewers are people who are just trying to have a good time on rare events. The emotion ranges change in frequent and infrequent viewers as there is bound to be more passion for the events in regular viewers, and the numbers reflect that. This crowd’s emotions and intentions directly impact the changes in the event’s performance. As the support and the cheering from the crowd audience act as the key encouragement in providing confidence oppositely, it can also make the supporting participant get demotivated and get distracted from the way to success [18]. That includes the performance low-quality performance of the participants. In the case of team performance in a sports event, the team members’ collaboration towards getting competitive advantages is crucial. The better version of one player in the team cannot make the whole team perform better. The under-confidence of the team members cannot make the team achieve more competitively. This study shows the advantages of computer-based analysis for event audience facial expressions.

**11. Conclusion.** It concludes all the reflections on the understanding of the facial emotions of the spectators from the live event. Then, they are decoded in the computer programming language, and an outcome is provided. It understands the emotions and sentiments of the people that affect the event’s environment. The final representation by the computer analysis makes the understanding of the changes provided by the spectators of the live event. The effect of the emotions and the audience’s behaviors in the crowd are computed using computer software analysis, and the effects of those reactions in the event are shown in this study.

It can be concluded that many factors are available related to the spectator’s emotions, behaviors, and involvement with life. The spectator’s involvement can be observed with their feelings and visible through their facial expression and body language. Computer vision and sentiment mainly provide all the information on a real-time basis and can be one reason for attending live sports. The spectator emotions and behaviors at live sports events using computer vision are increased in various factors. Fan engagement involves their feelings, visible through their facial expression and body language. Computer vision and sentiment mainly provide all

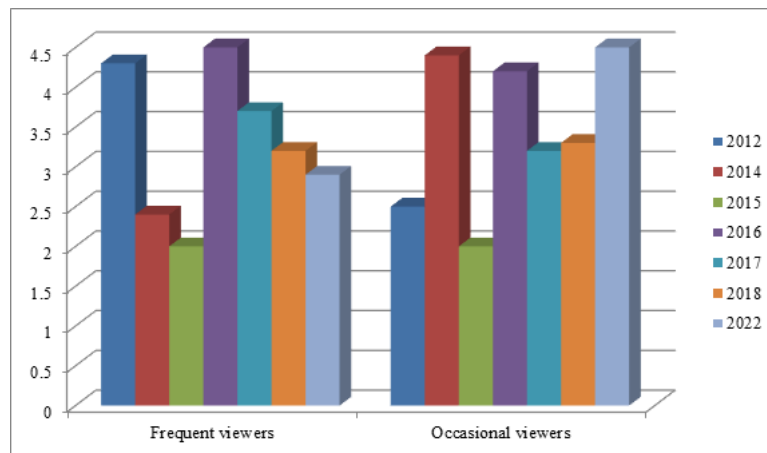


Fig. 10.2: Online and live viewership and emotion range statistics in recent years

the information on a real-time basis. Many people are available on the ground, and the management team looks after their safety and security. In the future, deep learning-based medical image diagnosis can be tested to improve performance.

#### REFERENCES

- [1] R. BISCAIA, M. YOSHIDA, AND Y. KIM, *Service quality and its effects on consumer outcomes: a meta-analytic review in spectator sport*, *European Sport Management Quarterly*, 23 (2023), pp. 897–921.
- [2] L. J. CABEZA-RAMÍREZ, F. J. FUENTES-GARCIA, AND G. A. MUÑOZ-FERNANDEZ, *Exploring the emerging domain of research on video game live streaming in web of science: State of the art, changes and trends*, *International Journal of Environmental Research and Public Health*, 18 (2021), p. 2917.
- [3] L. CAPASA, K. ZULAUF, AND R. WAGNER, *Virtual reality experience of mega sports events: A technology acceptance study*, *Journal of Theoretical and Applied Electronic Commerce Research*, 17 (2022), pp. 686–703.
- [4] Z. CHEN, Q. YANG, J. SHAN, T. LIN, J. BEYER, H. XIA, AND H. PFISTER, *iball: Augmenting basketball videos with gaze-moderated embedded visualizations*, in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 2023, pp. 1–18.
- [5] R. L. DOS SANTOS, M. D. L. M. PETROLL, R. BOEING, AND F. SCUSSEL, *Let's play a new game: the drivers of esports consumption*, *Research, Society and Development*, 10 (2021), pp. e40710515188–e40710515188.
- [6] M. FAN, A. BILLINGS, X. ZHU, AND P. YU, *Twitter-based birying: Big data analysis of english national team fans during the 2018 fifa world cup*, *Communication & Sport*, 8 (2020), pp. 317–345.
- [7] H. GONG, N. M. WATANABE, B. P. SOEBBING, M. T. BROWN, AND M. S. NAGEL, *Do consumer perceptions of tanking impact attendance at national basketball association games? a sentiment analysis approach*, *Journal of Sport Management*, 35 (2021), pp. 254–265.
- [8] C. HALKIOPOULOS, E. GKINTONI, H. ANTONOPOULOU, AND L. SKOUROLIAKOS, *Behavioral analysis of personality, branding and emotional state in e-sports*, *Technium Soc. Sci. J.*, 24 (2021), p. 434.
- [9] E. HEROLD, F. BORONCZYK, AND C. BREUER, *Professional clubs as platforms in multi-sided markets in times of covid-19: the role of spectators and atmosphere in live football*, *Sustainability*, 13 (2021), p. 2312.
- [10] S. KUMARESH, M. SHAH, AND S. SATHYANARAYAN, *analysis of opinionated texts on ipl 2020 matches using social media data*, *PalArch's Journal of Archaeology of Egypt/Egyptology*, 17 (2020), pp. 12564–12577.
- [11] Y. MAMO, Y. SU, AND D. P. ANDREW, *The transformative impact of big data applications in sport marketing: Current and future directions*, *International Journal of Sports Marketing and Sponsorship*, 23 (2022), pp. 594–611.
- [12] E. MENCARINI, A. RAPP, L. TIRABENI, AND M. ZANCANARO, *Designing wearable systems for sports: a review of trends and opportunities in human-computer interaction*, *IEEE Transactions on Human-Machine Systems*, 49 (2019), pp. 314–325.
- [13] Q. QIU, Y. ZUO, AND M. ZHANG, *Can live streaming save the tourism industry from a pandemic? a study of social media*, *ISPRS International Journal of Geo-Information*, 10 (2021), p. 595.
- [14] A. K. RAJENDRAN AND S. C. SETHURAMAN, *A survey on yogic posture recognition*, *IEEE Access*, 11 (2023), pp. 11183–11223.
- [15] M. SHARIATI FEIZABADI AND N. MOUSAVI, *Modelling cyber-aggression of basketball spectators (dematel technique)*, *Journal of Exercise and Health Science*, 1 (2021), pp. 35–52.
- [16] M. SJÖBLOM, J. MACEY, AND J. HAMARI, *Digital athletics in analogue stadiums: Comparing gratifications for engagement between live attendance and online esports spectating*, *Internet Research*, 30 (2020), pp. 713–735.

- [17] T. WANG AND Z. DU, *The impact of online sports watching atmosphere on viewers' continuous watching intention*, (2022).
- [18] B. H. YIM, K. K. BYON, T. A. BAKER, AND J. J. ZHANG, *Identifying critical factors in sport consumption decision making of millennial sport fans: mixed-methods approach*, *European Sport Management Quarterly*, 21 (2021), pp. 484–503.

*Edited by:* Sathishkumar V E

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

*Received:* Jun 10, 2023

*Accepted:* Aug 22, 2023



## SCALABILITY AND SUSTAINABILITY IN THE CONSTRUCTION OF A SOCIAL SPORTS MANAGEMENT INFORMATION PLATFORM BASED ON WEB TECHNOLOGY

**Abstract.** Scalability and Sustainability are a major traits that a website requires. Moreover, in Social sports management, information platforms must be reliable. Thus, different tools for developing a platform based on web technology are discussed in an empirical study for the purposes mentioned earlier. For the study, secondary data was analyzed, and qualitative methods were used. In addition, it was noticed that there are some problems related to a web technology-based platform. It was found that an improved development strategy in the beginning aids in traffic management. In addition, an improved tech stack is directly related to the data analysis process of a sports management system. A lack of such factors in the construction process reduced the scalability and sustainability of a social sports management information platform. Thus, a complete discussion aimed at understanding the sustainability and scalability of a web management platform is done.

**Key words:** Web technology, components of Web technology, frontend development, Backend development, Scalable website

**1. Introduction.** The incorporation of technology in a web-based platform decreases the hassle of looking for information. In addition, including a quality data management system helps manage all the essential aspects of sports management. Furthermore, including data modeling in a web-based platform ensures the reliability of the information [1]. Managing events through the analysis of historical data is one of the significant aspects of such a platform. However, the sustainability of web-based platforms might be an issue as a huge amount of data is processed in such models. Therefore, the following empirical study aims to understand and sustain a Social Sports Management Information Platform based on web technology [2]. In addition, the scalability of a web-based model is one of the primary concerns of the following study.

Figure 1.1 of the analysis related to the Stability of a scalpel website. During the analysis of past literature, it was noticed that including different analytical models in websites increases the load on a web platform [3]. Therefore, such construction of a web technology-based information system faces an issue of reduced sustainability. Moreover, handling a vast amount of data is compromised due to significant traffic. Thus, the construction of a web technology-based model is analyzed according to the collected data. In addition, all the necessary factors are analyzed for a better perspective of constructing a web technology-based platform. Further inclusion or relatable objectives in the study are a backbone for the empirical study [9]. Following the objections, a qualitative analysis was possible that helped to conduct systemic research. Including a problem statement is there, that helped to reflect an overall knowledge of the study's findings.

Web technology refers to the tools, technologies, and protocols used for building and accessing web-based applications and services over the Internet. It encompasses many components that enable the creation, delivery, and interaction of web content. Here are the main technical details that define web technology:

**HyperText Markup Language (HTML):** HTML is the standard markup language used to create the structure and content of web pages. It defines the elements, layout, and formatting of web content, including text, images, links, and multimedia.

**Cascading Style Sheets (CSS):** CSS is a stylesheet language used to control the presentation and styling of HTML elements. It allows designers to specify how the content should be displayed, including layout, colors, fonts, and responsive design for different devices.

**JavaScript:** JavaScript is a powerful scripting language in web browsers. It adds interactivity and dynamic behavior to web pages, enabling actions like form validation, animations, and real-time updates without page reloading.

**Web Servers:** Web servers handle HTTP requests and serve web pages to clients (usually web browsers). Popular web server software includes Apache, Nginx, and Microsoft IIS. These servers host the web application files and deliver them to users.

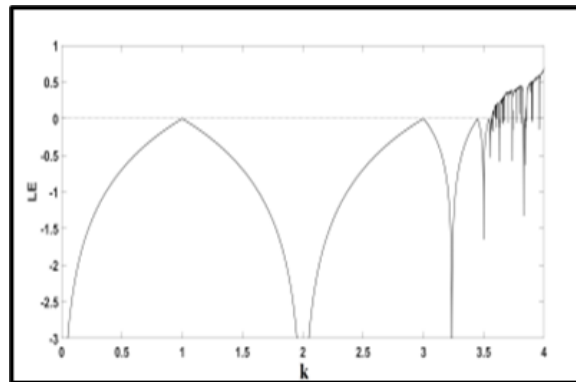


Fig. 1.1: Graph Explaining Stability In a Website

**HTTP (Hypertext Transfer Protocol):** HTTP is the fundamental protocol used for communication between web clients (such as browsers) and web servers. It defines how requests and responses should be structured, allowing the exchange of HTML documents, images, scripts, and other resources.

**Backend Development:** Backend technologies handle server-side logic, data processing, and database interaction. Common backend languages include JavaScript (Node.js), Python (Django, Flask), Ruby (Ruby on Rails), Java (Spring Boot), and PHP. Backend frameworks streamline the development of server-side functionality.

**Database Management:** Databases store and manage structured data web applications use. SQL (Structured Query Language) databases like MySQL, PostgreSQL, and NoSQL databases like MongoDB are commonly used. Databases allow for efficient data retrieval, storage, and management.

**APIs (Application Programming Interfaces):** APIs enable different software components (including web applications) to communicate with each other. Web APIs, often built using REST (Representational State Transfer) principles, allow developers to access specific functionalities or data from external services.

**Web Application Frameworks:** These software frameworks provide a structured way to build web applications. They often include tools, libraries, and pre-built components to simplify everyday tasks, enhance security, and improve efficiency.

**Web Security:** Web technologies incorporate security measures to protect users' data and ensure the integrity of web applications. Techniques include HTTPS (secure version of HTTP), encryption, authentication, and authorization mechanisms.

**2. Objectives.** To develop an empirical study in a systemic manner, having a pre-determined path is essential. Thus, the following objectives are created for studying the world as the backbone of the qualitative analysis.

1. To understand the working process of web technology-based platforms;
2. To understand the factors that are essential for web technology-based platforms;
3. To discuss the algorithm-based model to understand the proceeding of sports management;
4. To observe the factors impacting the process of web technology-based construction;
5. To elaborate on the problems of web technology-based platforms that are impacting sustainability and scalability.

**3. Methodology.** The development of a study methodology plays an important part. Moreover, the methodology of a study undertakes all the steps essential for reaching a conclusive result [11]. Therefore, for the constructive and systemic development of the study, the qualitative method of research was considered. Furthermore, secondary data was considered during the process of data collection. Secondary data is reliable and a pre-verified piece of information [8]. Hence, with the inclusion of secondary data for the research, the researcher was able to produce a reliable study with authentic data. In addition, quantitative analysis was



most reasonable for developing a conclusive outcome of the analysis. The qualitative process of analyzing data looks into different relations between factors [14]. In addition, qualitative analysis aids the process of reducing unrelated factors. Hence, for the reasons above, qualitative data was considered for the empirical study.

In a social sports management information platform based on web technology, several existing problems can arise that impact its effectiveness and efficiency. These problems may include:

1. **User Experience and Interface Design:** The platform's user interface and experience might not be intuitive or user-friendly, leading to difficulty navigating, accessing information, and performing tasks.
2. **Scalability Challenges:** As the platform gains popularity, it might struggle to handle a large number of users and data, leading to slow performance, crashes, or downtime during peak usage periods.
3. **Data Security and Privacy:** Protecting sensitive user data, such as personal information and payment details, from cyber threats and unauthorized access could be a significant challenge.
4. **Integration of Features:** Integrating various features like event scheduling, participant registration, payment processing, and performance tracking seamlessly can be complex and prone to glitches.
5. **Technical Compatibility:** The platform might not work well across different devices, browsers, or screen sizes, potentially limiting its accessibility for a diverse user base.
6. **Reliability and Uptime:** Ensuring the platform's availability 24/7 without disruptions or downtime is crucial for maintaining user trust and satisfaction.
7. **Content Management and Updates:** Managing and updating content, such as game schedules, news, and announcements, in a timely and accurate manner can be challenging.
8. **Community Engagement:** Encouraging active participation and engagement from users, such as players, coaches, and fans, can be difficult without effective communication channels and interactive features.
9. **Performance Analytics and Reporting:** Providing comprehensive analytics and reports on player performance, team statistics, and other relevant data might be inadequate or inaccurate.
10. **Payment and Financial Transactions:** Ensuring smooth and secure payment processing for event registrations, ticket sales, and other financial transactions could be problematic.
11. **Adaptability to Changes:** Rapid technological advancements might make the platform's technology obsolete or less effective over time, requiring constant updates and improvements.
12. **Regulatory Compliance:** The platform must adhere to legal and regulatory requirements related to data protection, online transactions, and other relevant laws.
13. **User Support and Assistance:** Offering timely and effective customer support to address user inquiries, troubleshoot issues, and provide assistance can be a challenge.
14. **Language and Cultural Diversity:** Serving a diverse user base with varying languages and cultural preferences may necessitate localization efforts.
15. **Monetization and Sustainability:** Developing a viable business model to sustain the platform's operations, cover costs, and generate revenue can be challenging.

**4. Components of a Web Technology-based Platform.** For a better building of a web technology-based social sports management information platform, there are specific components, which are necessary. Most of the features are based on the Java platform and provides different functionality to a web-based management system [5]. For instance, the dynamic behavior of a website is based on the integration of such components. For a web technology-based platform, database management is essential. In addition, for a reactive database system having such technology aids the sustainability and performance of the management system.

Table 4.1 of the study is related to different components related to the web-based modeling for a management system. All of the companies are related to database management and creating a webpage attractive and performing oriented [4]. Therefore, a systemic discussion of the aforementioned component in Table 4.1 is discussed in the following:

#### **A) Applet**

Applet is primarily a Java program that is associated with a Hypertext Mark-up Language (HTML) page. HTML and Cascading Style Sheets (CSS) aid the process of developing a web technology-based page [12]. However, Applet is one of the major components which provides added functionality to web pages

Figure 2 4.1 of the analysis is related to the lifecycle of Applet. It can be seen that in the initial phase,

Table 4.1: Different Components of a Web Technology-Based Platform

COMPONENTS	IDENTIFICATION	FUNCTIONALITY
<b>Applet</b>	A Java program in a hypertext mark-up language (HTML) page	Furnish different interactive features along with HTML alone
<b>JSP</b>	Java server pages	Provides a consistent way to extend web server functionality for dynamic web content
<b>Servlet</b>	Component architecture related to the Java programming model	Responds to client requests and develop dynamic responses
<b>EJB</b>	Enterprise java bean	Deliver portability across different servers
<b>Web services</b>	Modular applications	Works as an interface connecting provider and consumer

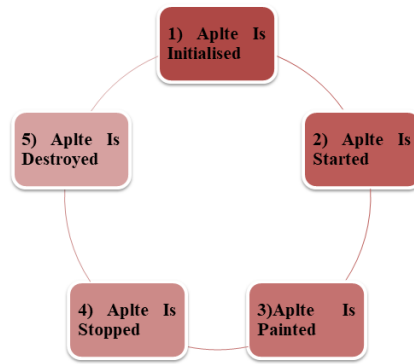


Fig. 4.1: Lifecycle of Applet

Applet is initialized, and after that Applet is started, and painted on a web page [6]. After the completion of the task, the Applet is stopped and destroyed. Hence, the above-mentioned lifecycle functionality of Applet is described.

**B) JSP**

Web technology-based management system containing JSP provides a consistent way to extend web server functionality for dynamic web content. Moreover, JSP provides reliability for a web server.

Figure 4.2 of the study describes the role of JSP in a database and web browsing interaction. It can be seen that JSP pages are developed with a technology that is not dependent on platforms and servers [13]. Therefore, element-based and scripting-based dynamic content can be included for better performance.

**C) Servlet**

In a dynamic web, technology-based model Servlet is a component architecture related to the Java programming model. The primary functionality of Servlet is, respond to client requests and develop dynamic responses.

Figure 4.3 of the empirical analysis provides a better understanding related to the working process of Servlet [10]. It can be seen that an HTTP request by a web browser is processed through a Servlet program and sent to the databases. Further, the Servlet program generates a response conveyed to the web browser.

**D) EJB**

An enterprise Java bean delivers portability across different servers. Moreover, EJB is the prime factor for maintaining the sustainability of a web page. Thus, for constructing a web technology-based sustainable and scalable social sports management information platform, EJB is an important component.

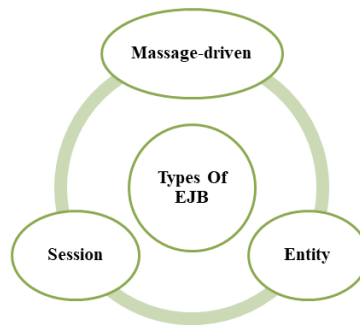


Fig. 4.2: Role of JSP

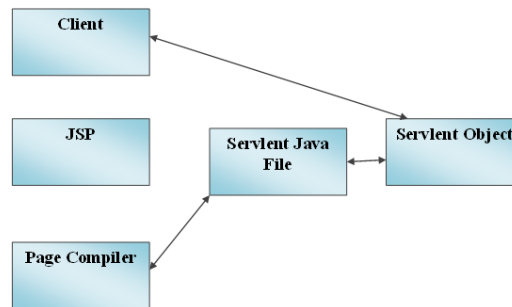


Fig. 4.3: Working process of Servlet

Figure 4.4 of the study discussed different types of EJB. It can be seen that there are session, entity, and message-driven EJB. The development of a social sports management information platform required information processing [15]. Thus, message-driven EJB is suitable for such a build.

#### ***E) Web services***

Most web services are Modular applications that work as a modulator interface, connecting provider and consumer.

Figure 4.5 discusses different languages used in web services available for the development of web-based platforms [7]. To make a social sports management information platform, choose an appropriate server depending on the planning process. Moreover, choosing a suitable server is required to be based on the aspect of scalability and sustainability.

**5. Process of Web Technology-based Construction.** Web technology primarily refers to using different tools and methods to establish communication between other devices. Moreover, all devices must contain the Internet to communicate [16]. In addition, having a database plays an essential role in developing a web technology-based sports management system. There are primarily two processes of development for a web technology-based management system. Front-end development related to the web page and back-end development contoured database management.

***A) Frontend Development:*** The front end of a website refers to the area with which the user immediately interacts. It is sometimes referred to as the application's "client side."

Table 5.1 of the study discusses different front-end languages used for developing a web page. Moreover, languages like HTML, CSS, and JavaScript are used to produce interactive web management [17].

Figure 5.1 of the empirical analysis represents a systemic amalgamation of different front-end languages. It can be seen that there are main languages used that help develop the front end of a web-based system. HTML provides a basic framework. In addition, CSS and its related tools help to make a page lucrative and interactive.

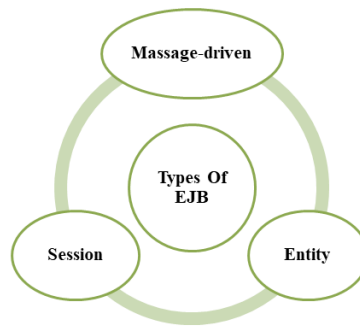


Fig. 4.4: Types of EJB

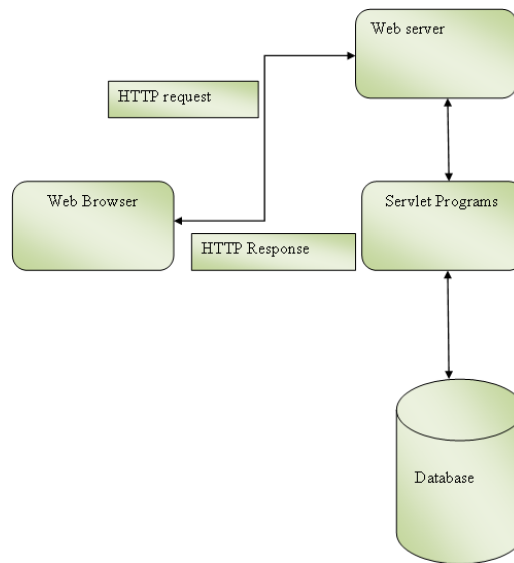


Fig. 4.5: Languages of web services

**B) Backend Development:** A website’s backend is its server side. The backend is the website area where clients or users cannot interact. In addition, a pseudo interaction is done between the backend and the user. In short, the Backend is the part of the website where data is stored in organized databases. Moreover, it is sometimes called the application’s ”client side.”

Table 5.2 of the study discusses different backend languages used in the backend management. In addition, features of the languages are discussed in the above table.

Figure 5.2 of the Study is related to the backend of a web technology-based portal. All of the languages gave different functionality, as mentioned in Table 5.2. However, most of the languages are based on framework management. In addition, interaction with data based is done by SQL and NoSQL. Additionally, package management is done with the languages [19].

Therefore, a sustainable and scalable web technology-based Sports management system is designed using such tools. In addition, the combination of different languages provides a robust structure to counter heavy traffic and make the management system interactive.

**6. Problem Statement.** During the analysis of the construction process of the web technology-based platform, it was noticed that sustainability is a significant issue. In addition, the scalability of such a platform

Table 5.1: Different front-end languages and their functionality

LANGUAGE	FEATURES
<b>HTML</b>	Hypertext Markup Language Hypertext is the linkage between web pages Markup is used to represent text documentation
<b>CSS</b>	Cascading Style Sheets Simplifies web design process Helps to make and change designs
<b>JavaScript</b>	Used for creating an interactive user experience
<b>AJAX</b>	Primarily communicate with servers without reloading a webpage

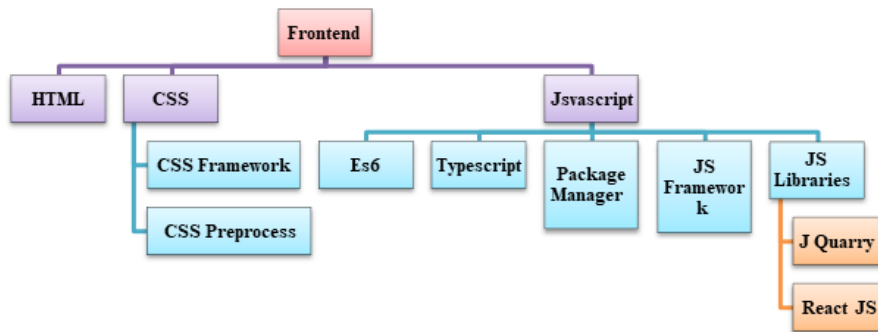


Fig. 5.1: Flow chart of Frontend

is one of the significant factors that impact the sustainability process. To make a Sclabael social sports management information platform, including different algorithms is essential. In. However, such inclusion has a significant impact on the working process of a web technology-based platform. Additionally, It was noticed that social sports management information platforms process a vast amount of data [16]. Therefore, the sustainability of such a platform is a huge concern.

Figure 6.1 of the empirical study illustrates a systematic representation of the problems in a social sports management information platform. During the construction of a Platform Based on Web Technology, certain factors impact the overall sustainability. For instance, a huge amount of data related to the design is the process of a website; further information decreases the performance of a web technology-based construction [18]. On the other hand, incorporating an analytical model for a social sports management information platform is required. Analytical models of a web technology-based platform introduce analytical features [19]. For instance, with such an analytical model, historical sports management data can be analyzed to make better predictions. However, including such data hinders a website’s sustainability and scalability. The reason for such hindrance is the intercity of an analytical model. Similarly, such websites process massive data sets [18]. In addition, increasing information is a significant problem with the scalability of a platform. Hence, the sustainability of a web technology-based platform is compromised due to the inclusion of considerable information.

Table 6.1 Discuss the problems in the designing process of a web technology-based platform. It can be seen that in order to counter the aforementioned problem designing a robust platform is important. The inclusion of such factors in the process of design aids in the process of managing traffic. Further factors like compatibility, tech stack, and performance of a website are mentioned in the table. It was noticed that all of the aforementioned factors have a proportional relationship with the sustainability and scalability of a web technology-based platform.

Table 5.2: Different backend languages and their functionality

Language	Features
PHP	Used for scripting to the server side
Node.js	Provides an environment for JavaScript codes
Python	System to uses robust database management
Ruby	Object-oriented programming language
Java	Used to make scalable web platforms
DBMS	Interacting with database
C#	Object-oriented programming

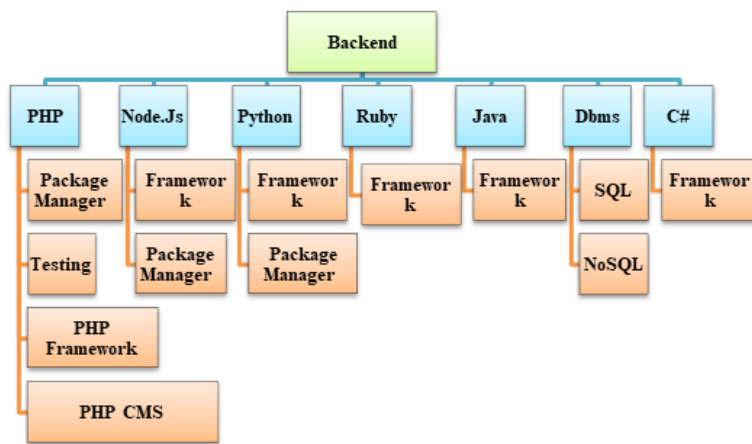


Fig. 5.2: Flow chart of Backend

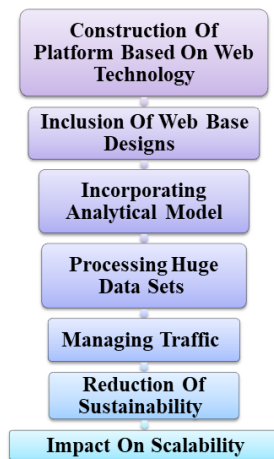


Fig. 6.1: Systematic representation of the possible problem in web technology-based platform

Table 6.1: Problems in the designing process

FACTORS	IMPORTANCE
<i>Development Strategy</i>	Aids the process of construction
<i>Tech Stack</i>	Improves the capabilities of web technology-based platforms thus supporting sustainability
<i>Compatibility</i>	Increases the accessibility of the platform, therefore, the sustainability of such a platform is supported
<i>Performance</i>	Aids traffic management, thus, the sustainability of a platform is supported,
<i>Unskilled Overlooking</i>	Impacts the sustainability of websites and hinders the debugging process
<i>Slow Servers</i>	Impacts the overall performance of a web technology-based design

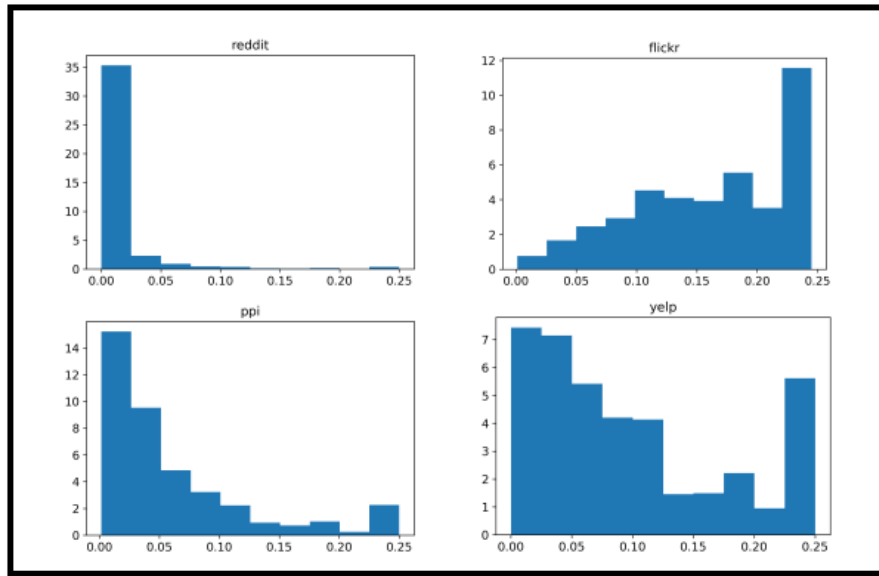


Fig. 7.1: Frequency distribution on diffusion weights over on individual dataset

**7. Result.** From the below figure of the dataset of diffusion weights, it can be said that the distribution of Reddit is 35 for the value of the standard deviation from 0 to the value 0.025. For the frequency of the distribution of Reddit, the value of the frequency distribution of the diffusion weights less than 5 for the other standard deviation. On the other hand, the value of the standard deviation for the flickr is quite distributive. All the value of the standard deviation is the highest from 0 to 0.25. The highest value of diffusion weights is 0.25, and the value is 11.5 (Frasca et al. 2020). The third value of the frequency distribution is the highest for the standard deviation of 0.00. The lowest value of the third distribution is for the standard deviation of 0.23. For the fourth frequency distribution, the highest value of the standard deviation is 0.02 and the lowest value is for the standard deviation of 0.23 (Frasca et al. 2020). Therefore, from the above graph, it is clear that the most frequent standard deviation is for the frequency distribution four and the least frequent standard deviation is for the frequency distribution first.

The above graphical analysis shows that the rate of the distribution of the frequency of the weights related to the diffusion of the individual dataset and the database. The lower distribution value of the first frequency diffusion weights is less effective for the growth and development of the sports management system which deals with the social and information platform (Frasca et al. 2020). Hence, the difference in the frequency distribution for the various datasets shows the variety in the data according to the scalability and the sustainability of the

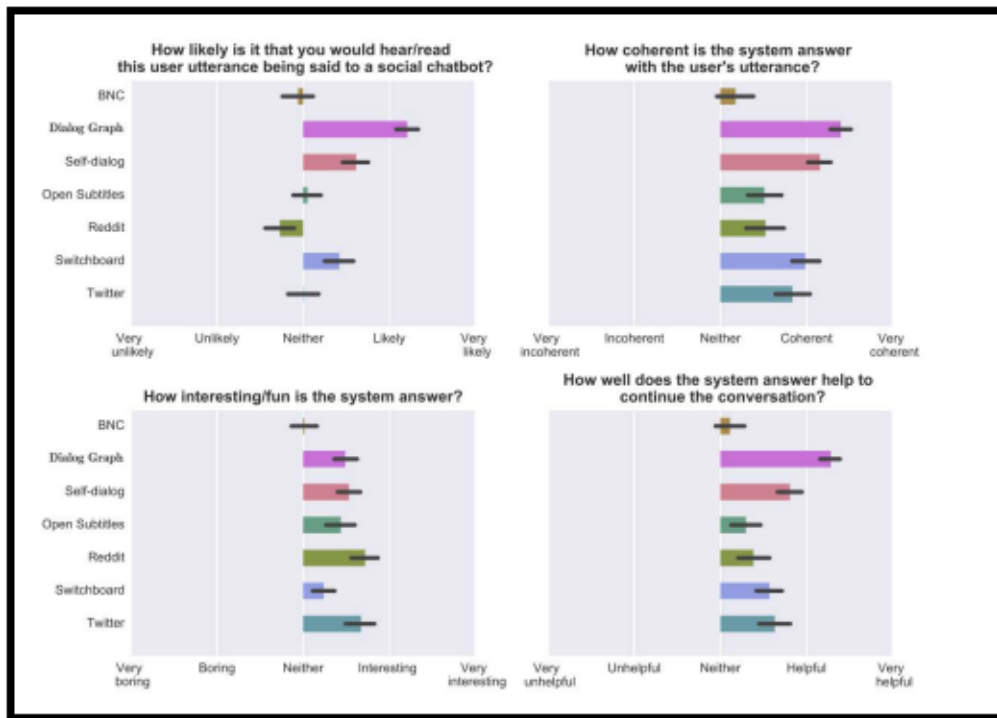


Fig. 7.2: 95% of the confidence intervals of the evaluated matrices

construction of the social sports management system.

From the above graph, it can be concluded that the use of the social chat board in the sports management system is the highest value for the dialog graph and most of the respondents like the dialog graph for the betterment of the sports authority. The least amount of the using app by the sports management system is the Twitter and the value of the usage of Twitter is nearly to the neither like or dislike portion (Jonell et al. 2019). Most of the people used to go to the coherent of the system in the sports management system and the highest number of the people who go for the coherent system is for the dialog graph. In the case of the interesting and fun part of the sports and the gaming system, most of the responders chose Twitter, and the app that used the least number is the BNC (Jonell et al. 2019). Therefore, the sports person and the sports management system have used the app Twitter at the highest number for advertisement of the game.

From the above graph of the confidence intervals of the evaluated matrices, it is clear that the conversation process between the sportsperson and the media is helpful for improvement and better performance. Also, the conversation between the peers of the sportsman and the other authority is constructive in increasing the rate of the version of the player (Jonell et al. 2019). Therefore, communication and conversation are of the most beneficial and effective processes to make sports management more powerful and effective for the growth and development of the sports authority.

**8. Conclusions.** Thus, a complete discussion of different web-based tools and technology is done in the study. A discussion related to different tools and technology helped to create an understanding of sustainable and scalable websites. It was understood that DBMS is a significant factor in developing a scabies sports management platform. In addition, a discussion of the problem is done that helps to understand the problem hence, in order to build a better social sports management information platform based on web technology, problems are required to be addressed. For the sustainability and scalability of such a web technology-based platform, a reduced load without compromising performance is important. Thus, addressing such issues with better technology is important. This study has extensively explored various web-based tools and technologies,



shedding light on the intricacies of creating sustainable and scalable websites. Notably, the significance of a robust Database Management System (DBMS) in the development of a social sports management platform has been emphasized. While this discussion has provided valuable insights, it's apparent that challenges remain on the path to crafting an enhanced social sports management information platform grounded in web technology.

To pave the way for the future, a proactive approach is imperative. The identified challenges and issues must be tackled head-on to ensure the seamless growth of the platform. Efforts should be directed towards refining the platform's architecture, optimizing its performance, and enhancing user experience. Striving for reduced load times while maintaining optimal performance levels will be key to sustaining scalability. This entails leveraging cutting-edge technologies that can handle increasing user demands without compromising the platform's efficiency. Furthermore, the integration of advanced data management techniques and cybersecurity measures will play a pivotal role in safeguarding user data and ensuring the platform's reliability. As technology evolves, continuous updates and adaptation will be necessary to keep pace with the ever-changing digital landscape.

In essence, this study provides a stepping stone for future advancements in the realm of social sports management information platforms. By addressing challenges and embracing emerging technologies, the future holds the promise of a more sophisticated, user-centric, and efficient platform that caters to the evolving needs of athletes, coaches, and enthusiasts alike.

#### REFERENCES

- [1] M. ABBAS, K. A. MEMON, A. A. JAMALI, S. MEMON, AND A. AHMED, *Multinomial naive bayes classification model for sentiment analysis*, IJCSNS Int. J. Comput. Sci. Netw. Secur, 19 (2019), p. 62.
- [2] G. ABEZA, N. O'REILLY, AND B. SEGUIN, *Social media in relationship marketing: The perspective of professional sport managers in the mlb, nba, nfl, and nhl*, Communication & Sport, 7 (2019), pp. 80–109.
- [3] L. ALMAZAYDEH, *Secure rgb image steganography based on modified lsb substitution*, International Journal of Embedded Systems, 12 (2020), pp. 453–457.
- [4] B. CHARBUTY AND A. ABDULAZEEZ, *Classification based on decision tree algorithm for machine learning*, Journal of Applied Science and Technology Trends, 2 (2021), pp. 20–28.
- [5] P. GOLPOUR, M. GHAYOUR-MOBARHAN, A. SAKI, H. ESMAILY, A. TAGHIPOUR, M. TAJFARD, H. GHAZIZADEH, M. MOOHEBATI, AND G. A. FERNS, *Comparison of support vector machine, naïve bayes and logistic regression for assessing the necessity for coronary angiography*, International journal of environmental research and public health, 17 (2020), p. 6449.
- [6] Z. HUANG AND Y. LIANG, *Research of data mining and web technology in university discipline construction decision support system based on mvc model*, Library Hi Tech, 38 (2019), pp. 610–624.
- [7] V. JACKINS, S. VIMAL, M. KALIAPPAN, AND M. Y. LEE, *Ai-based smart prediction of clinical disease using random forest classifier and naive bayes*, The Journal of Supercomputing, 77 (2021), pp. 5198–5219.
- [8] R. JOHANSSON, *On case study methodology*, Open house international, 32 (2007), pp. 48–54.
- [9] F. LORD, D. B. PYNE, M. WELVAERT, AND J. K. MARA, *Methods of performance analysis in team invasion sports: A systematic review*, Journal of sports sciences, 38 (2020), pp. 2338–2349.
- [10] J. MATEO-FORNÉS, A. PAGÈS-BERNAUS, L. M. PLÀ-ARAGONÉS, J. P. CASTELLS-GASIA, AND D. BABOT-GASPA, *An internet of things platform based on microservices and cloud paradigms for livestock*, Sensors, 21 (2021), p. 5949.
- [11] C. B. MEYER, *A case in case study methodology*, Field methods, 13 (2001), pp. 329–352.
- [12] L. MOYSIS, A. TUTUEVA, K. CHRISTOS, AND D. BUTUSOV, *A chaos based pseudo-random bit generator using multiple digits comparison*, Chaos Theory and Applications, 2 (2020), pp. 58–68.
- [13] A. PATEL AND S. JAIN, *Present and future of semantic web technologies: a research statement*, International Journal of Computers and Applications, 43 (2021), pp. 413–422.
- [14] M. PATEL AND N. PATEL, *Exploring research methodology*, International Journal of Research and Review, 6 (2019), pp. 48–55.
- [15] N. SALMI AND Z. RUSTAM, *Naïve bayes classifier models for predicting the colon cancer*, in IOP conference series: materials science and engineering, vol. 546, IOP Publishing, 2019, p. 052068.
- [16] A. SAWHNEY, M. RILEY, J. IRIZARRY, AND M. RILEY, *Construction 4.0*, Sawhney, A., Riley, M., Irizarry, J., Eds, (2020).
- [17] S. M. SEPASGOZAR, *Digital twin and web-based virtual gaming technologies for online education: A case of construction management and engineering*, Applied Sciences, 10 (2020), p. 4678.
- [18] D. TITH, J.-S. LEE, H. SUZUKI, W. WIJESUNDARA, N. TAIRA, T. OBI, AND N. OHYAMA, *Application of blockchain to maintaining patient records in electronic health record for enhanced privacy, scalability, and availability*, Healthcare informatics research, 26 (2020), pp. 3–12.
- [19] I. ZHOU, I. MAKHDOOM, N. SHARIATI, M. A. RAZA, R. KESHAVARZ, J. LIPMAN, M. ABOLHASAN, AND A. JAMALIPOUR, *Internet of things 2.0: Concepts, applications, and future directions*, IEEE Access, 9 (2021), pp. 70961–71012.

*Edited by:* Sathishkumar V E

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

*Received:* Jun 30, 2023

*Accepted:* Aug 22, 2023



## REINFORCEMENT LEARNING-BASED ALGORITHMS FOR MUSIC IMPROVISATION AND ARRANGEMENT IN SENSOR NETWORKS FOR THE INTERNET OF THINGS

**Abstract.** The process of learning any new technology requires acquiring the best knowledge about the information of that technology. The better the knowledge humans get about digital technology, the more they become efficient in implementing technological development. In developing the musical rhythm and tuning, the application of programming technologies helps improve the quality. In constructing networking sites and sensing technologies, algorithmic learning processes help in effective development. This development occurs by making the systematic process of transforming a data processing language and data interpreter. Thus, it helps in performing programming effectively in the present as well as future purposes. Therefore, it reflects all the benefits of machine learning. Thus, the preference for machine learning increases technological impact. This development of the programming used in the computer makes humans learn about something easily and get the best information. The effectiveness of the technological development by the algorithm used in the data processing implements the best way to improve the technological language transformation from human language to computer operating language. There is a transnational perspective of the average beat commonness of each part of the music. “Reinforcement algorithms-based learning” incorporated with sensor networks has proposed compelling opportunities for improving “music improvisation” and interpretation.

**Key words:** Machine Learning, Learning-Based Algorithms, Music Improvisation, Sensor Networks, Internet of Things

**1. Introduction.** In the learning system of the world, the influence of technology has greatly increased with its development and easy accessibility. The technological implementation helps in learning new things quickly and efficiently. In technological development, introducing technology based on learning algorithms greatly improves students. This is the case of the development of music learning, which also helps set up digital music technologies. It became a preference of people in 2010 about 50 percent. This technology includes the utilization of networking sites and software-based technologies for the transformation of digital data. The learning-based algorithms help develop the programming used in the computer for making humans learn about something quickly and get the best information about that topic. The improvisation in musical technologies helps in creating a better musical rhythm. On the other hand, the application of the sensor networks collects information from various locations and controls the collected data. This monitoring of the development in the technologies of musical development makes the learning process easily understandable. This also makes the computer programming used for the learning algorithm easy to understand.

**1.1. Objectives.** Some basic objectives are correctly identified and described in this study. This includes the basic concept of the technology used in the making of the development of computer data programming and the development of sensor technologies. The improvisation of music between the computer networking systems and the digital networking is processed by the application of learning technologies of algorithms. This includes the processing of different types of sensors of digital technology and the collection of information from additional monetized data for improving communication in computer operations. This helps in the development of the operating surroundings of the computer user in the making of improvisation of data processing more easily and rapidly done. Some of the objectives of the multi-networking technologies of sensor networking and music improvisation in the connection between computer technologies and humans are as follows:

1. To elaborate on the concept of algorithm-based learning
2. To describe the importance of algorithm-based learning in musical development
3. To identify the issues generated in the handling of sensor networks in IoT
4. To provide some of the network programming development by the application of algorithmic programs
5. To examine the impact of machine learning in algorithmic language
6. To state some of the factors in machine-based learning used in the sensing networks of the musical development.

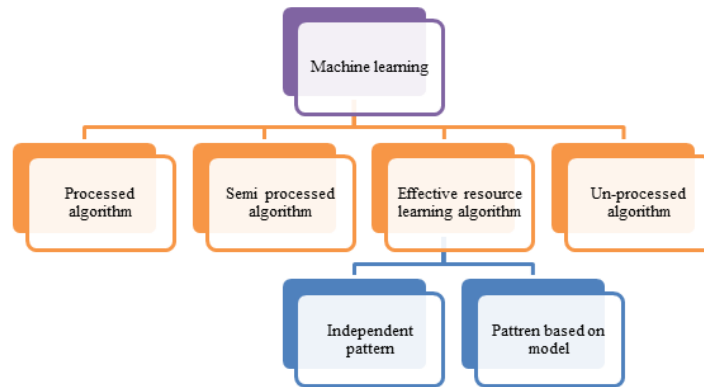


Fig. 2.1: Algorithm-based Machine learning

The main contribution of the work is:

1. The research underscores the importance of acquiring in-depth knowledge about digital technology for efficient implementation. It emphasizes that an enhanced understanding of technology translates into better utilization and implementation of technological advancements.
2. The research highlights the role of programming technologies in improving the quality of musical rhythm and tuning. By applying programming techniques, the research contributes to the advancement of music production and enhances the overall musical experience.
3. The study showcases the significance of algorithmic learning in developing networking sites and sensing technologies. This contribution elucidates how algorithmic processes aid in effectively creating and enhancing these technologies.

**2. Proposed Methodology.** In data collection of machine-based learning and improvisation, the computer technologies used in the improvisation of musical development are collected from the descriptive investigation. This includes collecting information from the newspaper, articles, and journals based on machine-based algorithmic languages. It represents all the information about networking technologies used to collect information from various locations. This collection process includes the following digital data patterns in calculating data processing. This processing makes the mathematical computation by comparing all the collected information from the network station. That information helps in constructing a better and more effective way in the implementation of the musical data, including the rhythmic structure of the music.

**2.1. Concept of Algorithm-based Learning.** Learning by machine-based programs helps in providing adequate information to the learner. This helps in the collection of information based on the topic. Information collection is based on networking programs used in the operating system. This learning program includes the functioning of algorithmic commands inputted by the user in the learning programming [5]. The inputted data instructs the operating system to process the data and get the best output by comparing all the related information. This information sets or classifies the information into various types namely processed data, semi-processed and unprocessed data as shown in Figure 2.1. This represents the mathematical calculation of the information based on the reliability of the information changed as per time.

This information shown in Table 2.1 gives the best knowledge provided by the operating computer program to humans. The comparison of the data changes as per the changes of time and the necessity of the programmer. The categories of the classification of the data include the following independent pattern to process the inputted data and the model-based pattern [4]. Moreover, all knowledge algorithms are commonly connected in their ability to represent information from processed data and utilize those learning to make future estimations

Table 2.1: Application benefits of algorithm-based learning

Process of algorithmic learning	The function of algorithmic learning
Processed algorithm	Application of logical regression for obtaining numerous data for learning
Semi-processed algorithm	Semi-processed learning is used for setting connections between IV
Effective resource learning algorithm	Classifying output data for making effective decision
Un-processed algorithm	Comparing all related information about the inputted information Presenting valuable information for making effective decision making

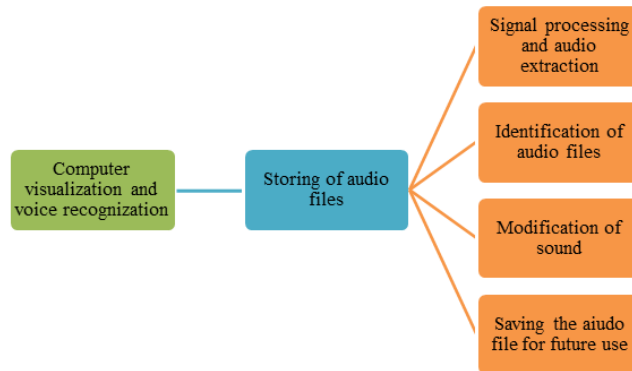


Fig. 2.2: Algorithm-based learning in musical development

concerning new input.

**2.2. Importance of Algorithm-based Learning in Musical Development.** The application of the algorithm in the composition makes the utilization of various data by comparing the necessity of the application. This is the programming of techniques that helps in the utilization of the programming in the computation process as shown in Figure 2.2. The composition of the music through computer programming requires the setting up of data comparison between the application of algorithmic process in the making of the music and the stored data in the data storage [6]. This computer programming helps in the making of the development in music by dividing it into various types. This classification creates the effective decision-making capability of the digital computer-based music composer. One of the best ways of making the development of music composition by the algorithm is the examination and composition of music in various programming.

Algorithm-based learning plays a crucial role in the development of music, revolutionizing the way music is created, produced, and experienced. The importance of algorithm-based learning in musical development can be outlined as follows:

1. **Enhanced Creativity:** Algorithms provide new avenues for musical creativity. By leveraging algorithms, musicians and composers can explore unique patterns, harmonies, and melodies that might not have been intuitive through traditional methods. This leads to the creation of novel and innovative musical compositions.
2. **Automated Composition:** Algorithm-based learning enables the automation of certain aspects of musical composition. Algorithms can generate musical sequences, chord progressions, and motifs, freeing composers to focus on more intricate and expressive elements of their work.
3. **Pattern Recognition and Analysis:** Algorithms excel at recognizing patterns and trends within

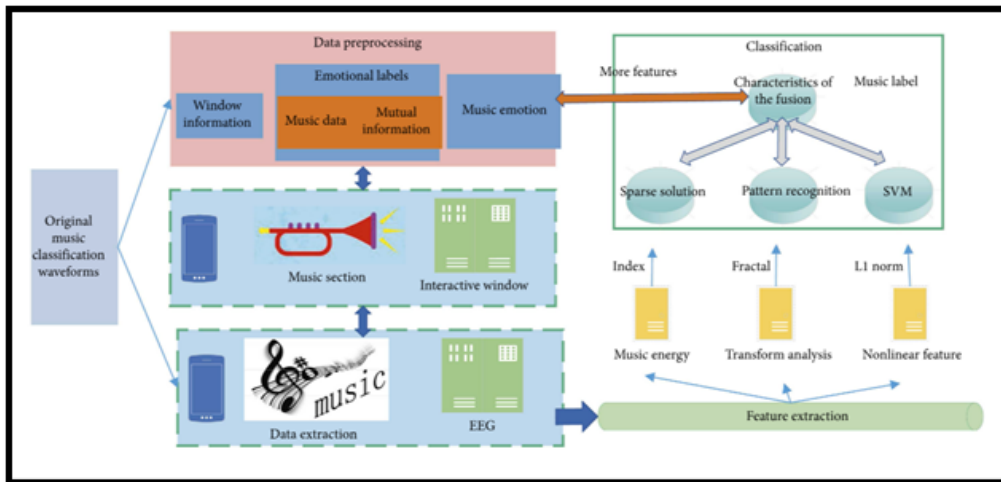


Fig. 2.3: Application process of machine learning in Music development

large datasets. In the context of music, algorithms can analyze vast amounts of musical pieces to identify common motifs, chord progressions, and structures. This insight aids musicians in understanding established conventions and experimenting with variations.

4. **Personalized Musical Experiences:** Algorithmic learning facilitates the creation of personalized musical experiences for listeners. Music streaming platforms use algorithms to recommend songs and playlists based on listeners' preferences, leading to tailored and enjoyable musical journeys.
5. **Real-Time Performance and Interaction:** Algorithm-based learning has enabled real-time musical performance and interaction. Musicians can use algorithms to respond dynamically to their performance, creating live electronic music and interactive installations that respond to audience engagement.
6. **Music Production and Sound Design:** Algorithms contribute to enhancing music production and sound design. From generating unique soundscapes to automating mixing and mastering processes, algorithms streamline production workflows and improve the quality of the final product.
7. **Genre Fusion and Experimentation:** Algorithms can blend elements from different musical genres and styles, fostering genre fusion and experimentation. Musicians can explore uncharted territories by combining diverse musical elements in innovative ways.
8. **Efficient Learning and Teaching Tools:** Algorithm-based learning offers efficient tools for learning and teaching music. Interactive apps and platforms utilize algorithms to provide instant feedback on technique, composition, and musical theory, accelerating the learning process.
9. **Reviving Historical Styles:** Algorithms can analyze historical compositions and recreate musical styles from different eras. This enables musicians to study and reintroduce classical and traditional styles in contemporary contexts.
10. **Collaboration Between Humans and Machines:** Algorithm-based learning fosters collaboration between human musicians and machines. Musicians can use algorithms to generate initial ideas, which can then be refined and personalized with human creativity and emotion.

In summary, algorithm-based learning has transformed musical development by expanding creative horizons, automating certain processes, providing personalized experiences, and enabling dynamic real-time interactions. It empowers musicians, composers, producers, and listeners alike to explore new dimensions of music and contribute to the evolution of musical artistry.

This programming development includes the collection of information based on rhythmic choices such as MIDI for the structuring of the music lyrics. Then the processed music in setting the music sound enhances

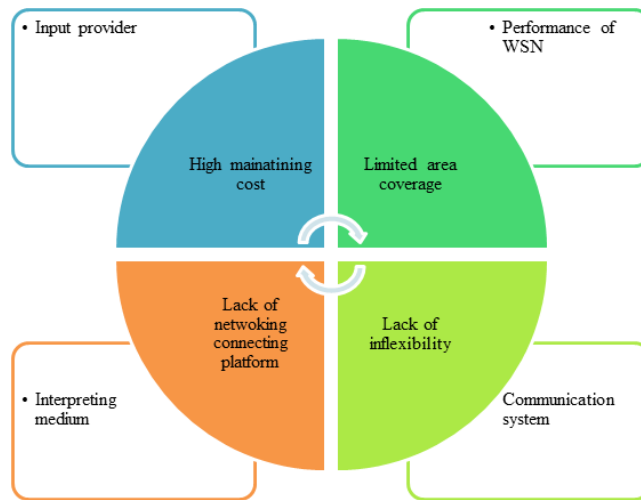


Fig. 2.4: Ges in the handling of sensor networks in IOT

Table 2.2: Challenges of WSN in IOT

Categories of challenges	Description
Models of transformation Models used in mathematical calculation Lack of information based on the system Grammatical structure Execution view Developing methods Crossbreed systems	These challenges in the wireless networking sites increase the issue of data storing safety by the storing network. This mainly depends on the lack of a networking platform and the communication system used in the process of wireless networking sites of the internet

the quality of the sound. This development in the quality and the rhythmic way of the music depends on the program used in computer programming [3]. The application of algorithms in the current days is popularly known as **Auto tune**. This basically enhances the quality and tuning of the music by computer programming languages.

**2.3. Challenges in the Handling of Sensor Networks in IoT.** Wireless networking technologies include the utilization of networking sites and software-based technologies for the transformation of digital data. The learning-based algorithms help develop the programming used in the computer for making humans learn about something easily and get the best information. Although it has some major challenges that deflect the improvisation of computer programming as shown in Figure 2.4. These challenges in the wireless networking sites increase the issue of data storing safety by the storing network [1]. This mainly depends on the lack of a networking platform and the communication system used in the process of wireless networking sites of the internet. This challenge makes the performance of data transferring through wireless networks deteriorate.

This challenge affects the obstruction of the generated risk in the wireless networking of the Internet of Things. This also affects the transformation medium making the smooth running of the programming languages as shown in Table 2.2. This depends on the decision-making capability of the digital computer-based music composer [9]. One of the best ways of developing music composition by the algorithm is the examination and composition of music in various programming.

**2.4. Network Programming Development by the Application of Algorithmic Programmes.** Network programming development through the application of algorithmic programs has significantly transformed

the landscape of networking and communication technologies. This convergence of algorithmic prowess with networking has yielded a range of benefits and advancements:

1. Algorithmic programs streamline the processing and transmission of data across networks. They optimize data routing, compression, and encryption, ensuring efficient data exchange between devices.
2. Algorithms enable dynamic routing and load balancing in network traffic. Adaptive routing algorithms distribute data packets effectively, optimizing network utilization and minimizing congestion.
3. Algorithmic programs play a crucial role in network security. They underpin intrusion detection systems, firewall configurations, and encryption protocols, safeguarding data and systems from cyber threats.
4. Algorithms ensure optimal QoS by prioritizing critical data traffic over less time-sensitive data. This is particularly essential for applications requiring low latency, such as video conferencing and online gaming.
5. Algorithmic programs contribute to the development of network protocols and standards. They dictate the rules for data transmission, error correction, and synchronization, ensuring interoperability across diverse devices.
6. Algorithms automate network monitoring and management tasks. They analyze network performance metrics, detect anomalies, and trigger actions for fault management and performance enhancement.
7. Algorithmic programs enable real-time data analytics within networks. They process incoming data streams to derive meaningful insights, supporting decision-making and predictive analytics.
8. SDN leverages algorithms to dynamically manage network traffic flows and configurations. This programmability enhances network flexibility and agility, adapting to changing requirements.
9. Algorithms facilitate network virtualization and cloud computing. They allocate resources efficiently, manage virtual network instances, and optimize data placement across cloud servers.
10. Algorithmic programs enable seamless connectivity in the Internet of Things (IoT). They ensure efficient communication between numerous IoT devices, managing data exchange and synchronization.
11. Algorithmic programs drive emerging technologies like edge computing and 5G networks. They optimize data processing at the network edge and facilitate ultra-low-latency communication.
12. Algorithms contribute to network optimization by minimizing latency, maximizing throughput, and optimizing resource allocation. This enhances the overall performance and user experience.
13. Algorithms enable predictive maintenance by analyzing network behavior patterns and anticipating potential failures. This proactive approach minimizes downtime and enhances network reliability.
14. Algorithmic programs provide scalability and flexibility to networks. They adapt to changing demands and scale resources dynamically to accommodate varying workloads.

The system of network interpreting technologies helps in the making of development through the application of communication. This includes the making of the development in interpreting technology such as regular monitoring and system development. This interface creates a better medium for accessing the programming languages [8]. Thus the programming in the procedures develops the performance of different operating systems performance. The description of the database servers and the systematic processes make the performance of the work by comparing all the circumstances of the programming. The system uses one server as the client server and others for communicating with different databases. This process helps in interfacing multiple users in one server without disrupting any of the users.

**2.5. Examine the Impact of Machine Learning in Algorithmic Language.** The application of digital technologies helps in the development of the programming so the applicants get the benefits of using the programmed language easily. The application of the technologies includes the application of the system in daily life application. This includes the implementation of different technologies in the process of language transformation of the developments including the development in the quality and the rhythmic way of the music depending on the program used in computer programming [7]. The application of algorithms in the current days is popularly known as **Auto tune**. This basically enhances the quality and tuning of the music by computer programming languages. It also includes features like the **recognition of voice** and programming of the software application. The development of networking sensors makes the development of the transforming process used in the continuation of data transformation from different locations easy. These procedures of



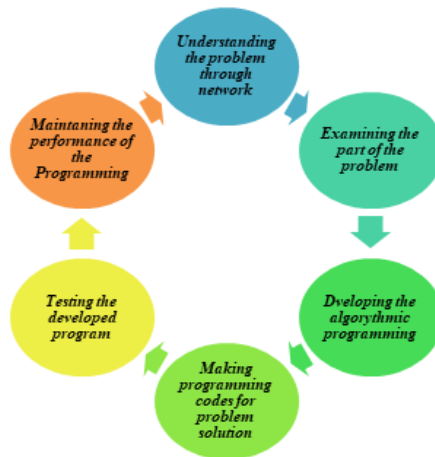


Fig. 2.5: Network programming development by the application of algorithmic programs

Table 2.3: Network programming development by the application of algorithmic programs

Programming and network development factors	Impacts of Programming and network development factors.
Understanding the problem through a network Examining the part of the problem Developing the algorithmic programming Making programming codes for problem solution Testing the developed program Maintaining the performance of the Programming	The application of the sensor networks collects information from various locations and also controls the collected data. This monitoring of the development in the technologies of musical development makes the learning process easily understandable. This also makes the computer programming used for the learning algorithm easy to understand.

machine learning make the effective understanding of the best and easier way of transforming the value.

As shown in Figure 2.7 it also makes the database servers and the systematic processes help the performance of the work by comparing all the circumstances of the programming. The system uses one server as the client server and others for communicating with different databases. This process helps in interfacing multiple users in one server without disrupting any of the users [10]. Therefore the making of development through the applications of machine technologies makes the development of humans become the most effective technology performer. In the networking platform and the communication system used in the process of wireless networking sites of the internet.

**2.6. Factors in Machine-based Learning used in the Sensing Networks of the Musical Development.** The application of machine-based learning increases the development of performing any task through business applications. In network development machine learning helps in constructing a better way of utilizing program software [2].

This also depends on the better functioning of some of the factors in the developing process such as:

*Improvement of training time* Training time is very important in network data transformation. This makes the development of the learning time of the information from digital platforms. This helps in making different technologies in the process of language transformation. This includes the application of AI in the working procedures of network programming.

*Improvement of testing time.* The development of software programming helps in the development of computer programming. This helps in making the development of the application process of the software

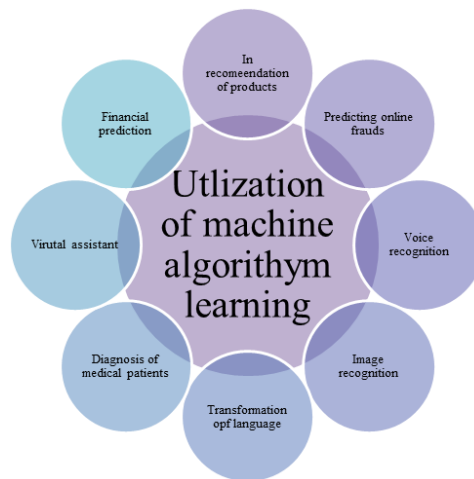


Fig. 2.6: Examine the Impact of Machine Learning In Algorithmic Language

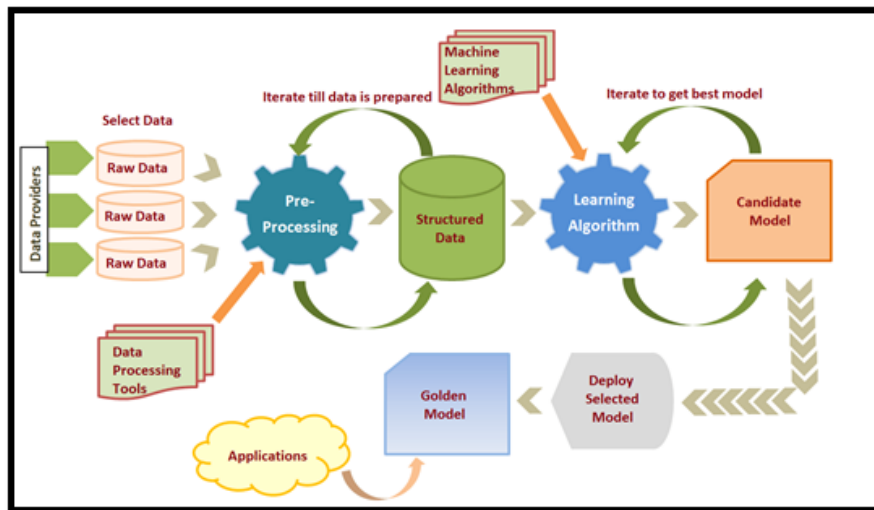


Fig. 2.7: Processes of Machine Learning In Algorithmic Language

Table 2.4: Impact of machine learning on algorithmic language

Utilization of algorithm learning in programming development	Impacts of the utilized techniques in computer programming
In recommendation of products Predicting online frauds Voice recognition Image recognition Transformation of language Diagnosis of medical patients Virtual Assistant Financial prediction	The application of the technologies includes the application of the system in daily life application. This includes the implementation of different technologies in the process of language transformation. The development of networking sensors makes the development of the transforming process used in the continuation Procedures of machine learning make the effective understanding of the best and easier way of transforming the value.

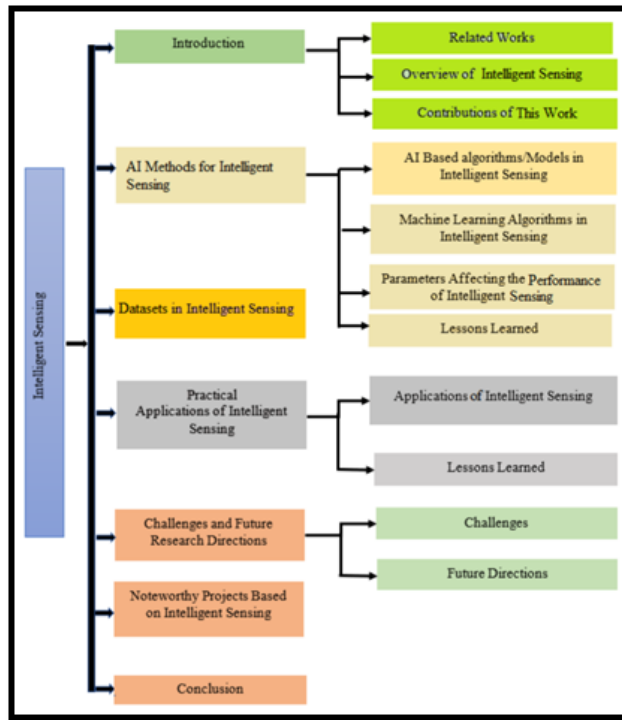


Fig. 2.8: Process of machine-based learning used in the sensing networks of the musical development

process perform quickly. Thus the programmer gets results in less time.

*Accuracy of the program.* The performance of the programming technology improves the ability to represent information from processed data and utilize that learning to make future estimations concerning new input. Thus the efficiency of the utilization of programs improves.

**2.7. Problem Statement.** The application of machine-based learning makes the introduction of the best utilization of computer programs. This program reflects the input of the challenges in the wireless networking sites increasing the issue of data storing safety by the storing network. This mainly depends on the lack of a networking platform and the communication system used in the process of wireless networking sites of the internet [3]. This challenge makes the performance of data transferring through wireless a network deteriorates. In the development of music, rhythm helps in making the development of composting technology. This application of machine learning enhances the quality and tuning of the music by computer programming languages [2]. This problem in the storing of the network of wireless network and communicating with the computer OS was unable to reflect the fluctuating issues generated every day of the programming. This representation helps to provide all the information about the computer-based learning process by the algorithm to state the effective impact of machine-based learning in the internet mediums.

**3. Result Analysis.** The below figure depicts a transnational perspective of the average rhythm commonness of each part of the music. Whenever it corresponded to the residuals rigidly, the rhythm-specific practice acquainted on 8 trials does not exceed the note-specific practice acquainted on 24 trials [11]. However, this outcome does not mean that the growth of the training “set size”, that is the “rhythm frequency” is less than 3, which is produced by 24/3. The above figure has been described as a transnational perspective of the average rhythm commonness of each part of the music, where this numeral is greater than 3 for all pieces. This result has displayed that “rhythm context” is instructive for graphic elements, particularly, explicit timing [11]. However, notations of the same “rhythm context” definitely do not pursue precisely the same way as lyrical

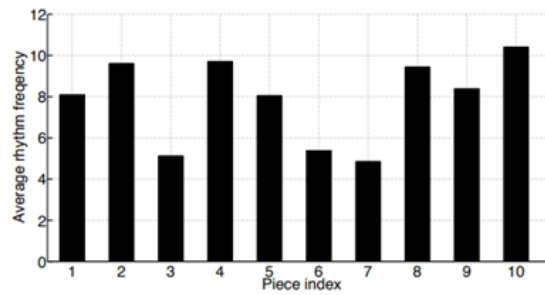


Fig. 3.1: A transnational perspective of the average rhythm commonness of each part of music

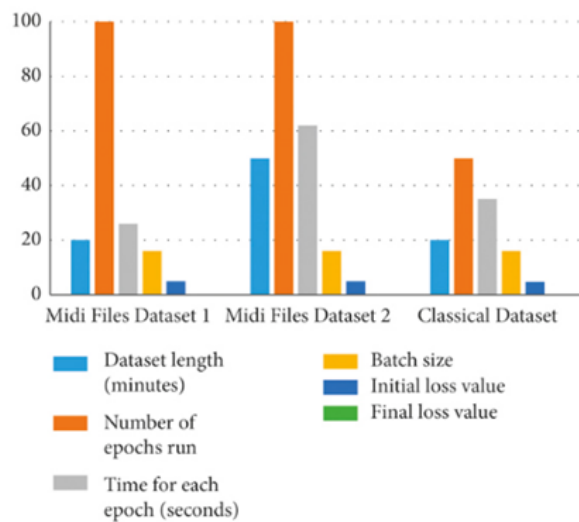


Fig. 3.2: Suggested model outcome for various datasets

phrases. Simply utilizing the “rhythm context” is not sufficient; in order to gain a more acceptable prognosis and additionally decrease the activity “set size”.

Figure 3.1 has articulated suggested model outcomes for various datasets. The data sets are “midi files dataset 1”, “midi files dataset 2”, and the “classical dataset” [12]. It is done by evaluating the dataset size in minutes, the numeral of ages vary, the period for each age in seconds, set size, initial losing value, and last losing worth. It has been found that “reinforcement learning-based algorithms” incorporated with sensor networks propose compelling opportunities for improving “music improvisation” and interpretation. Intelligent agents can comprehend and adapt their musical determination, resulting in more graphic and interactive music concerts in the IoT era, by leveraging real-time data from detectors [12].

Figure 3.2 has also depicted that sensor networks played an integral function in catching and transferring data that are related to music. Various sensors can be employed, such as “microphones for audio data”, “biosensors for catching physiological indications” or “accelerometers for motion data”. Valid sensor arrangement, data accompaniment, and bluster removal processes are essential for sufficient music data assembling.

Figure 3.3 has depicted the suggested model outcomes for the failure and age of various datasets. The datasets are “midi files dataset 1”, “midi files dataset 2”, and the “classical dataset” [12]. It is obvious from the picture that as the numeral of age is improved the failure worth declines. The failure worth at the 10 age is

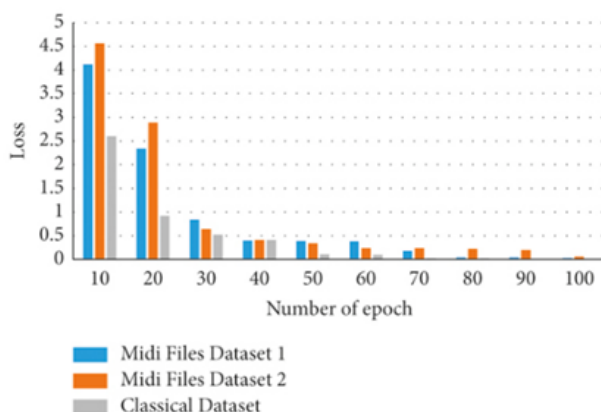


Fig. 3.3: Suggested model outcomes for the failure and age of various datasets

4.12 for “midi files dataset 1”, 4.56 for “midi files dataset 2”, and 2.60 for the “classical dataset”. In the middle of the age when the worth of these datasets is 50, the loss value is 0.39 for “midi files dataset 1”, 0.34 for “midi files dataset 2”, and 0.11 for the “classical dataset” [12]. The failure worth at the 100 age is 0.03 for “midi files dataset 1”, 0.06 for “midi files dataset 2”, and 0.01 for the “classical dataset”.

The result has also shown that “incorporating reinforcement learning algorithms” into “sensor networks” allowed “real-time decision-making” based on data accumulation. By incorporating the strength of “reinforcement learning” with “sensor data”, agents can acclimate their improvement and engagement techniques based on the modifying “musical context” seized by the sensors. This integration has improved the responsiveness and interactivity of music concerts.

**4. Conclusion.** In applying an algorithm-based learning process for Music Improvisation and structuring of Sensor Networks for the Internet of Things. This technique of computer-based programming helps in the learning-based algorithms for developing the programming used in the computer. This is used to make humans learn about something quickly and get the best information about that unaware information. Including improvisation in musical technologies helps in creating a better musical rhythm. Also, the application of the sensor networks collects information from various locations and controls the collected data. The process of monitoring the development of musical development through programming technologies makes the learning process easily understandable.

#### REFERENCES

- [1] S. A. A. ABIR, A. ANWAR, J. CHOI, AND A. KAYES, *Iot-enabled smart energy grid: Applications and challenges*, IEEE access, 9 (2021), pp. 50961–50981.
- [2] H. ALI, I. K. NIAZI, B. K. RUSSELL, C. CROFTS, S. MADANIAN, AND D. WHITE, *Review of time domain electronic medical record taxonomies in the application of machine learning*, Electronics, 12 (2023), p. 554.
- [3] R. K. BEHARA AND A. K. SAHA, *Artificial intelligence methodologies in smart grid-integrated doubly fed induction generator design optimization and reliability assessment: A review*, Energies, 15 (2022), p. 7164.
- [4] B. GOWRISHANKAR AND U. B. NAGAPPA, *Automated transcription for raga recognition and classification in indian classical music using machine learning*, in 3rd International Conference on Integrated Intelligent Computing Communication & Security (ICIIC 2021), Atlantis Press, 2021, pp. 211–217.
- [5] S. KUMAR, K. SINGH, S. KUMAR, O. KAIWARTYA, Y. CAO, AND H. ZHOU, *Delimitated anti jammer scheme for internet of vehicle: Machine learning based security approach*, IEEE Access, 7 (2019), pp. 113311–113323.
- [6] H. MA AND X. PANG, *Research and analysis of sport medical data processing algorithms based on deep learning and internet of things*, IEEE Access, 7 (2019), pp. 118839–118849.
- [7] T. MAHMOOD, J. LI, Y. PEI, F. AKHTAR, A. IMRAN, AND K. U. REHMAN, *A brief survey on breast cancer diagnostic with deep learning schemes using multi-image modalities*, IEEE Access, 8 (2020), pp. 165779–165809.

- [8] P. MANICKAM, S. A. MARIAPPAN, S. M. MURUGESAN, S. HANSDA, A. KAUSHIK, R. SHINDE, AND S. THIPPERUDRASWAMY, *Artificial intelligence (ai) and internet of medical things (iomt) assisted biomedical systems for intelligent healthcare*, *Biosensors*, 12 (2022), p. 562.
- [9] L. MOYSIS, L. A. ILIADIS, S. P. SOTIROUDIS, A. D. BOURSIANIS, M. S. PAPADOPOULOU, K.-I. D. KOKKINIDIS, C. VOLOS, P. SARIGIANNIDIS, S. NIKOLAIDIS, AND S. K. GOUDOS, *Music deep learning: Deep learning methods for music signal processing-a review of the state-of-the-art*, *IEEE Access*, (2023).
- [10] M. VAEZI, K. NOROOZI, T. D. TODD, D. ZHAO, G. KARAKOSTAS, H. WU, AND X. SHEN, *Digital twins from a networking perspective*, *IEEE Internet of Things Journal*, 9 (2022), pp. 23525–23544.
- [11] G. G. XIA, *Expressive collaborative music performance via machine learning*, CMU Ph. D. thesis, (2016).
- [12] P. S. YADAV, S. KHAN, Y. V. SINGH, P. GARG, R. S. SINGH, ET AL., *A lightweight deep learning-based approach for jazz music generation in midi format*, *Computational Intelligence and Neuroscience* (2022).

*Edited by:* Sathishkumar V E

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

*Received:* Jul 3, 2023

*Accepted:* Aug 22, 2023



## VULNERABILITY DETECTION IN COMPUTER NETWORKS USING VIRTUAL REALITY TECHNOLOGY

**Abstract.** This paper challenges the time-related challenges inherent in conventional network security detection methodologies. It is achieved by incorporating virtual reality technology into the domain of computer network security detection. The research methodology employs optimization calculations to extract attributes that characterize network security vulnerabilities. Concurrently, the weighting of diverse vulnerability attributes is adjusted using a web crawler, a comprehensive list of injection points, and meticulous analyses of the attacks' genetic characteristics. This collective approach facilitates the exploration of automated network security vulnerability detection within a virtual reality framework. The study's empirical results demonstrate that the detection method proposed within this investigation exhibits a notably reduced delay of 75.33 milliseconds. The respective delays observed in the two conventional methods stand at 290.11 milliseconds and 337.30 milliseconds. The substantial decrease in detection delay validates the effectiveness and efficiency of the devised automated network vulnerability detection approach grounded in virtual reality technology.

**Key words:** Network security vulnerabilities, Virtual reality technology, Detection methods, Optimization calculations, Automated detection

**1. Introduction.** The abnormal network traffic attacks are evolving into more complex and diversified. This anomalous traffic can disrupt the regular functioning of terminal PC systems at a minor scale, while at a critical level, it can induce server immobility and network breakdowns. Consequently, automation of network vulnerability detection serves a dual purpose: upholding the network's robust operation and aiding network administrators in preemptively identifying vulnerabilities for proactive security measures.

Automated tools for detecting security vulnerabilities prove highly effective in safeguarding network integrity. These tools systematically scan the network for vulnerabilities, quickly identifying and generating reports on potential weaknesses. The utilization of such tools substantively elevates network security and curtails the susceptibility to network breaches. Throughout the automated vulnerability detection process, the system treats each node vulnerability as an independent entity, readjusting the vulnerability's weight attribute differentially through calculation and optimization to derive an attack graph. This graph represents the potential for unauthorized entities to penetrate and attack the network system via vulnerabilities, showcasing their varying weight inequalities. The construction of an attack graph entails the aggregation of vulnerability clusters. With the evolution of network technology, the difficulty of network security challenges is simultaneously escalating, encapsulating various issues like network attacks, data leaks, malware incursions, and phishing exploits [19].

Integrating virtual reality technology into network security management involves various domains such as enterprise network platforms, computer programming, and the global network. This integration has generated an excess of research outcomes. For instance, within enterprise network platforms, authentication virtual reality technology is employed to curtail user system access actions. An illustrative instance involves merging fingerprint identification technology with network access rights management, thus transposing the customer's tangible identity into the virtual domain for multifaceted authentication of their distinct physical identity. When challenging network security issues restricting from malicious network viruses, remedies involve utilizing network encryption virtual reality technology and critical management virtual reality technology to counteract these threats [13]. Among the pivotal strategies for network security management lies security vulnerability detection. The convergence of network security vulnerability detection and virtual reality technology remains relatively

limited in the research scope. Nevertheless, considering the efficacy of applying virtual reality technology to network security management, amalgamating these two realms will emerge as a pivotal avenue for future field advancement.

The paper is structured as follows: Section 2 investigates the literature review, providing an overview of existing approaches, their limitations, and the need for more efficient solutions. Section 3 presents the proposed method, detailing the integration of virtual reality technology, optimization calculations, web crawlers, and simulation of attacks for automatic detection. Section 4 presents the results and discussion, showcasing experimental outcomes that validate the effectiveness of the proposed method compared to traditional approaches, highlighting reductions in detection time delays. Finally, Section 5 summarizes the findings in conclusion, emphasizing the potential of virtual reality technology to reshape the landscape of network security vulnerability detection, improving efficiency and minimizing risks.

**2. Literature Review.** Information security vulnerabilities encompass imperceptible conditions that can compromise or harm information systems' confidentiality, integrity, and availability. Users engaging with information can construct autonomous environments through networking, leveraging high-performance computing systems to execute computational tasks from diverse locations. This approach moderates the burden of system management, boosting utilization efficiency. However, inherent security vulnerabilities exist within the virtualization software domain, posing threats that can severely undermine information services' dependability on virtualized environments.

Systems grounded in virtual technology exhibit complexity, and safeguarding information resources proves to be a complex undertaking. Lapses in caution could culminate in the emergence of information security vulnerabilities. Even when comprehensive measures are undertaken to strengthen and support the entire system's information security to the highest level, there remains the certainty that the level of protection and prevention might fall short, thereby exposing information to potential risks [7, 5].

Network attacks stand out as one of the most prevalent network security challenges. Hackers may exploit vulnerabilities, weak passwords, social engineering, and other techniques to infiltrate systems, appropriate, sensitive information, manipulate data, or disrupt operations. The data leakage issue is equally concerning, which can expose users' details like names, addresses, social security numbers, and credit card information. The malware hazard further intensifies network security, encompassing viruses, trojans, worms, system damage, data theft, or surveillance of user activities. Additionally, phishing represents a deception wherein fraudsters manipulate users into exposing personal information or login credentials through email, SMS, and social media.

The network environment represents high openness, sharing, and interactivity, catering to diverse user operational requirements. However, this expansive range of network functionalities and services simultaneously introduces numerous vulnerabilities, imperiling user information security. As computer network runtime increases, the quantity and intricacy of exposed security vulnerabilities tend to magnify. Among these, link connection vulnerabilities arise during network information interactions. The link serves as a pivotal conduit for transmitting computer network information. If subjected to malicious attacks, it could result in issues like information destruction, loss, and network security incidents.

Typical methods for network vulnerability detection involve constructing a sample set model and deriving network variables for simulation experiments. An example is the application of hidden Markov models to shape the network information sample set, thus facilitating the automated refreshment of network security vulnerability detection. However, this approach's attaining the network flow table is intricate, leading to diminished detection efficiency and presenting an efficiency difficulty [6, 1]. An automatic detection method is formulated by extracting dynamic numeric variables and scrutinizing static attack processes by fusing dynamic and static aspects. Notably, this method's assessment has been confined to simulations on open-source software. Consequently, the outcomes of these experiments are partial, casting uncertainty over their practical significance.

### 3. Proposed Methodology.

**3.1. Automatic detection method of network security vulnerabilities under virtual reality technology.** At the core of network security vulnerability, technology analyses the program's corresponding source code and identifies vulnerabilities. Numerous latent vulnerabilities exist within a program's source code, which malicious actors could exploit to breach systems, extract information, or disrupt operations.



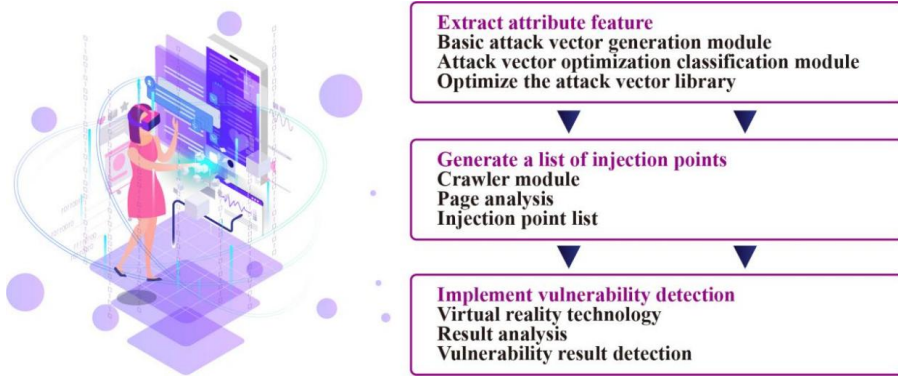


Fig. 3.1: Overall structure of the proposed method

Consequently, the scrutiny of program source code is of paramount importance in the pursuit of detecting network security vulnerabilities. In virtual reality technology, the automated identification of network security vulnerabilities hinges on using virtual reality to simulate and detect malicious attacks. This process automatically generates an attack plan based on the identified findings [17]. The schematic representation of the proposed method is illustrated in Figure 3.1.

While creating an attack graph, we emphasize achieving automatic construction. This entails crafting an attack vector repository via optimization calculations. The role of analyzing attack injection points encompasses pinpointing network injection sites, acquiring web pages through web crawlers, and subsequently subjecting the web pages to analysis via these crawler tools. The progression of attack and analysis involves emulating the attack sequence employing virtual reality technology. Subsequent scrutiny of the yielded outcomes facilitates the ultimate determination of vulnerabilities.

**3.2. Extraction of network vulnerability feature attributes.** Throughout the automated network security detection process, the network system perceives vulnerabilities within the nodes as if they were its code. Optimization calculations adjust the weight attributed to distinct vulnerability characteristics, giving rise to an attack diagram. Over time, attackers will exploit the network via diverse vulnerabilities, and the visible weight disparities between these vulnerabilities will become apparent [12]. The composition of an attack graph involves incorporating a vulnerability group technique. At a given point in time, denoted as “ $t$ ”, within a cluster featuring “ $m$ ” vulnerabilities, it is postulated that each vulnerability houses “ $k$ ” individual genes. The vulnerability gene studied in this paper adopts the vulnerability gene, exposure gene and repair gene, so  $k$  is 3, and the location of a vulnerability can be shown as  $q_i = (q_{i1}, q_{i2}, q_{i3})$ , where  $i = 1, 2, \dots, m$ . In the process of generating the attack map, the starting position of the vulnerability can be regarded as a point in the spatial coordinates, so the location of the vulnerability can be represented by digitizing when the coordinates are iterated. Set the movement speed of the vulnerability  $v_i = (v_{i1}, v_{i2}, v_{i3})$ ,  $v_{i1}, v_{i2}, v_{i3}$  are the movement speed of the vulnerability in the three coordinate axes, and its attribute can be expressed as:

$$V_{i+1} = \omega \vartheta_{ir} + \varepsilon_1 (z_{ir} - p_{ir}) + \varepsilon_2 (z_{ir} - p_{ir}) \quad (3.1)$$

In the above formula,  $\varepsilon_1$  is the learning coefficient,  $\varepsilon_2$  is the network learning coefficient,  $\omega$  is the inertia coefficient,  $\vartheta_{ir}$  is the initial speed of the vulnerability,  $z_{ir}$  is the optimal location of the vulnerability, and  $p_{ir}$  is the probability of the vulnerability being applied in the attack, then  $p_{ir}$  meets:

$$\sum_{i=1}^m p_{ir} = 1, \quad 0 < p_{ir} < 1 \quad (3.2)$$

When  $\omega$  is between 0 and 1, vulnerabilities in motion can converge, so when the value of  $\omega$  is determined, the number of malicious attacks can be predicted. When the vulnerability in the network is exposed to a large

Table 3.1: Analysis of characteristics and attributes of network security vulnerabilities

S.No.	Vulnerability characteristic attribute	Size	Number of source code files	Total lines of source code	Version number
1	Blender	144.6 MB	4126	1628412	2.78c
2	Clang	124.6 MB	2258	1324761	3.9.1
3	Crystal Space	489.9 MB	4421	944126	2.0
4	Firefox	1.65 GB	21274	7321054	50.0.1
5	MPlayer	116.3 MB	3215	4521585	5.2.2
6	Mysql	482 MB	5569	5445102	10.3.6
7	PHP	149.3 MB	1932	4568941	Va3.8
8	Nebula Device 2	86.6 MB	1106	2384161	4.1.3
9	OpenSceneGraph	166.4 MB	2601	1158523	4.1.0

extent, the degree of repair will also increase, and the vulnerability will be difficult to be applied. Therefore, the three attributes of vulnerability will increase with time, so  $p_{ir}$  will decrease, and the extraction of vulnerability feature attributes is completed.

**3.3. Generation of injection point list.** Upon utilizing the attributes extracted to represent vulnerability features, the subsequent step involves scrutinizing injection points, culminating in creating a roster of injection points. This undertaking necessitates the support of a web crawler [4, 9]. The web crawler essentially functions as a program that autonomously gathers these vulnerability feature attributes. Through analysis of these attributes, it determines the sequencing within the URL queue for capture while also retaining the pertinent attributes.

The web crawler can proficiently retrieve pages from the target system within vulnerability detection. This process involves identifying forms, input fields, and other elements and subsequently dissecting their properties. Based on this attribute information, vulnerability detection programs can identify and append reasonable vulnerability junctures to the injection point list. The collated attributes are categorized into nine groups, and the configuration for vulnerability analysis is outlined in Table 3.1.

Many redundant segments within the URL queue tend to emerge during the crawling process, impacting the crawling pace. In response, we have opted for implementing the BloomFilter algorithm to deduce, identify, adjust, and store vulnerabilities before inclusion in the URL queue procured through crawling. Once this determination is reached, the queues not belonging to the collection are incorporated, thus achieving the desired collection. An "ID" value of "Y" signifies the presence of this queue within the collection. Subsequently, the queue is introduced into the "urlset" to finalize the deduplication process, culminating in compiling the injection point list, outlined in Table 3.2.

The system employs the forms retrieved from the attack injection point for vulnerability analysis and documentation, generating a list of attribute injection points. Subsequently, these analyses and records facilitate an extensive exploration of vulnerabilities, culminating in the automated selection of the most suitable detection approach [15].

**3.4. Implementation of vulnerability detection under virtual reality technology.** In virtual attack testing, guided by the information encapsulated in the generated injection point list, virtual reality technology is harnessed to emulate attack behaviours and engage with the server. In the context of employing virtual reality technology to detect network security vulnerabilities, it proves beneficial for reproducing the network environment. This allows security experts to immerse themselves in the network environment, facilitating a firsthand comprehension of the location, nature, and repercussions of security vulnerabilities.

For instance, within a virtual reality setting, security experts possess the capacity to enact attack scenarios, assess the network environment's resilience against these attacks, and uncover and rectify latent vulnerabilities.

Table 3.2: List of attribute injection points

URL Queue Name	Source code file name	Starting line number of the duplicate part of the queue	Relative error in duplication (attack input)	Maximum observed relative error (corresponding input)
Blender	Key.c	668	9.12E-12 (2.06E12)	6.12E-11 5.66E14)
OpenScene	Shande out set.c	251	6.61E-11 (6.32E02)	3.54E-61 (1.42E03)
Device 2	Sc allouy.c	2456	6.51E-10 (1.11E12)	5.61E-03(1.55E06)
Blender	key.c	214	5.66E-09 (6.30E11)	7.31E-08 (1.15E10)
MPlayer	note_draw.c	1032	7.54E-12 (1.25E04)	8.36E-06 (6.58E03)
Crystal	audio_concica.c	825	5.32E-11 (8.36E06)	9.54E-13 (9.32E10)
Blender	key.c	10	4.33E-12 (9.34E08)	7.96E-01(6.47E04)
Clang	pixman_gration.c	251	6.48E-04 (6.97E11)	8.45E-11 (6.32E06)
Firefox	opus_Active.c	483	1.61E-09 (8.99E13)	9.44E-02 (7.35E11)
OpenScene	cairo_out.c	125	5.64E-11 (4.66E13)	6.21E-07 (9.32E12)
Firefox	sd_root.c	2156	8.34E-14 (6.32E05)	3.14E-02 (6.05E11)
Graph	cairo_strock.c	232	6.24E-09 (7.98E13)	4.20E-06 (6.21E01)
OpenScene	Shande out	18	1.98E-01 (9.63E07)	9.34E-05 (6.14E05)
PHP	Shortest_path_suppression.c	231	9.61E-81 (5.68E09)	9.01E-08 (6.77E12)

Throughout this process, due to the attacker's virtual identity, the attribute of network security vulnerabilities can be construed as the fundamental genetic element for constructing a virtual attack. The established virtual reality environment and virtual attacks are tools for extracting information embedded within network security vulnerabilities. Subsequently, the security vulnerability classification is executed by categorizing vulnerability, exposure, and repair genes in alignment with the classification of attack information [3]. This delineates the functional framework within the realm of vulnerability detection for virtual attacks.

In the detection process, after orchestrating artificial virtual vulnerability attacks through gene selection and the cross-assembly of virtual attacks, the results of these attacks are inputted into the calculation of a fitness function. This yields vulnerability information and other pertinent data, ultimately detecting network security vulnerabilities. The research on automatic detection methodologies for network security vulnerabilities under the purview of virtual reality technology has been concluded.

**4. Results and Discussion.** A simulation experiment using network security vulnerabilities is constructed within the virtual reality technology to validate the credibility of the automated detection approach presented in this paper. This analysis aims to compare the detection outcomes with those yielded by two conventional vulnerability methodologies, thus substantiating the merits of the proposed approach. The illustrative network structure for the experimentation is depicted in Figure 4.1. The figure illustrates the presence of five nodes: A, B, C, D, and E. Node A represents the external network intrusion host. Within the firewall, the four nodes have distinct roles: furnishing network services, managing database services, safeguarding and overseeing servers, and preserving critical files. It's worth noting that intranet external network access and intranet-to-intranet communication do not necessitate firewalls. Nevertheless, when an extranet node seeks access to an intranet node, passage through the firewall is mandatory, granting exclusively network server access [18, 11].

It is important to acknowledge that specific actual vulnerabilities within the intranet can potentially target nodes within the intranet. The vulnerability configurations for each intranet node are listed in Table 4.1.

Upon establishing the experimental environment, three distinct methods are employed for the detection: the conventional modelling detection method, the numerical detection method, and the network security vulnerability automatic detection method developed within the framework of virtual reality technology as outlined in

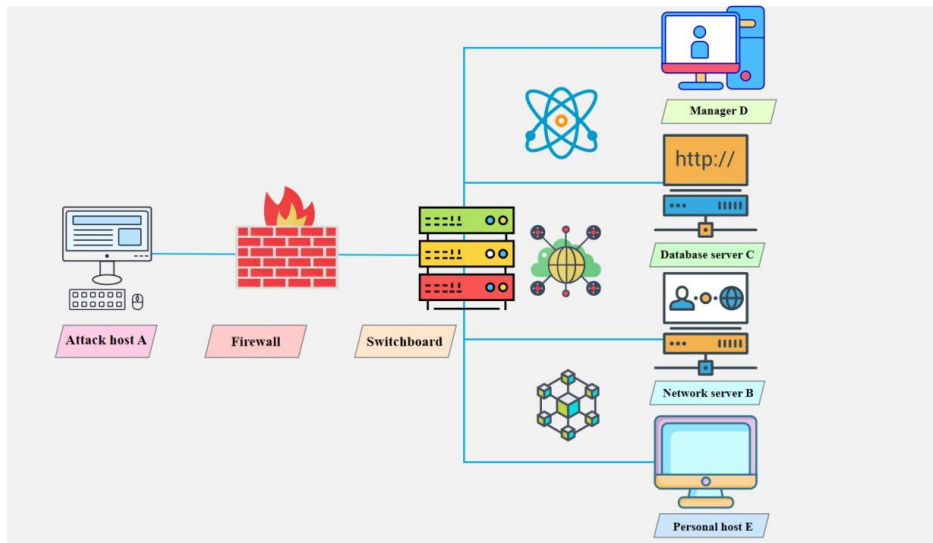


Fig. 4.1: Construction of network topology

Table 4.1: Intranet node vulnerability settings

S.No.	Node	Vulnerability number	Server or software	Privilege escalation	Ease of use (%)
1	Network server B	Mc10-062	www	U-R	70
2	Network server B	Mc11-002	www	U-A	100
3	Database server C	Mc11-012	Oracle	O-A	90
4	Database server C	Mc10-065	Oracle	O-U	85
5	Manager D	Mc10-045	Windows	O-A	100
6	Personal host E	Mc10-006	Office	O-A	70

this paper. The Figure 4.2 illustrates a structured listing of vulnerabilities linked to diverse nodes, servers, and software components within the network environment. Each entry is distinguished by a unique vulnerability number, denoting its specific characteristics. The corresponding server or software connected to each vulnerability is specified, alongside an assessment of the level of privilege escalation that the vulnerability permits.

Moreover, the exploitability of each vulnerability is represented as a percentage, shedding light on the relative complexity of exploiting each instance. The vulnerabilities span across a range of nodes, encompassing network servers, database servers, manager nodes, and personal hosts. The figure furnishes an in-depth understanding of the vulnerabilities infusing the network environment, explaining their relevant attributes and potential implications for security.

**4.1. Detection of interactive frequency.** To advance the detection process, a key measure is to elevate the frequency of network interactions. To enhance the precision of vulnerability detection outcomes, 70 detection samples are used to monitor the variations in interaction frequency among the three methods throughout the detection procedure. The time consumption for each method is illustrated in Figure 4.3.

The depicted figure highlights a discernible trend: as the quantity of test samples grows, the method introduced in this paper consistently exhibits an escalation in response requirements. During the initial 20-second detection window, the modeling detection method gains a notable edge, owing to its coverage of a portion

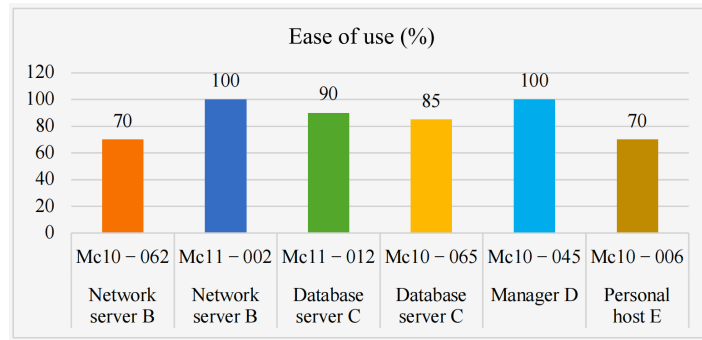


Fig. 4.2: Overview of vulnerabilities in the network environment

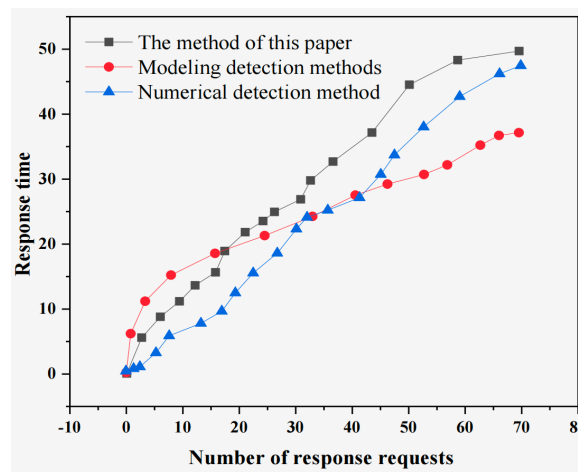


Fig. 4.3: Comparison of interaction of response time

of the sample data. As the testing duration reaches the 20-second mark, compared to the conventional two methods, the method outlined in this paper surpasses them regarding response requirements within the same time span. This interaction frequency results in a more incredible count of detected vulnerabilities, showcasing distinct advantages [10, 2].

**4.2. Detection of vulnerabilities.** As established in the preceding article, the simulation network vulnerability detection experiment is conducted for the four designated target nodes - B, C, D, and E. During this experiment, the time delay for vulnerability detection is logged across the three methods and subsequent detection outcomes are compared. This information is outlined in Table 4.2.

The data presented in the table above clearly indicates that the detection delay associated with the method detailed in this paper stands at 75.33 milliseconds. This represents a significant reduction of 290.11 milliseconds and 337.30 milliseconds compared to the respective delay times of the two traditional methods. This achievement not only translates to minimized detection delays but also substantiates the efficacy of the designed automated detection approach for network security vulnerabilities within the context of virtual reality technology [16, 8]. Virtual reality technology systems are commonly employed in addressing network security challenges. Constructed upon virtual reality technology, a novel scheme for automated network security vulnerability detection emerges, effectively improving historical issues of continued delays and detection efficiency in vulnerability assessment. Furthermore, this scheme strengthens the network security vulnerability detection capacity.

Table 4.2: Comparison of test results of three methods

Method	Target node	Number of detected vulnerabilities	Detection delay (ms)
The method of this paper	B	Mc10-062	75.33
		Mc11-002	
	C	Mc11-012	
		Mc10-065	
		Mc10-045	
Modeling detection method	B	Mc10-062	365.44
	C	Mc10-065	
	D	Mc10-045	
	E	-	
Numerical detection method	B	Mc10-062	412.63
		Mc11-002	
	C	-	
	D	Mc10-045	
	E	Mc10-006	

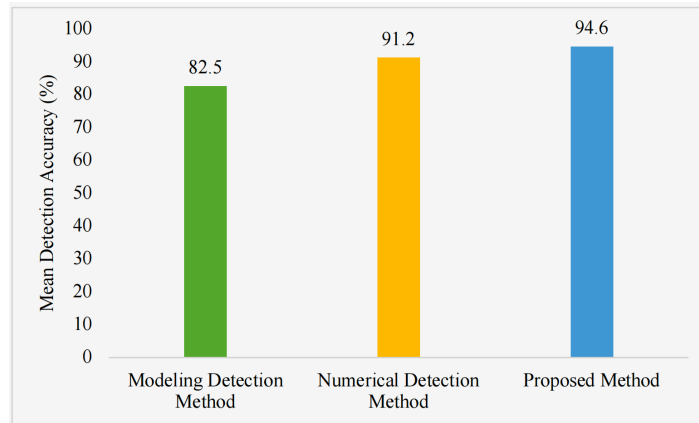


Fig. 4.4: Comparison of mean accuracy detection

A comparison of the three detection methods varying mean detection accuracy in vulnerability identification is shown in Figure 4.4. The modeling detection method achieved 82.5% accuracy, displaying proficiency in recognizing pattern-based vulnerabilities, yet showing potential limitations against anomalies and zero-day vulnerabilities. Outperforming this, the numerical detection method reached 91.2%, employing advanced algorithms for heightened precision in detecting known vulnerabilities and refined deviations. However, the proposed method excelled significantly with 94.6% accuracy, capitalizing on virtual reality and optimization to demonstrate exceptional adaptability, versatility, and effectiveness in identifying vulnerabilities, particularly within complex scenarios and emerging threats.

Virtual reality technology plays a dual role in this context. Firstly, it replicates authentic network environments encompassing network topology, servers, and network devices. This gives security experts deeper insights into network operations and potential vulnerability locations. Through the lens of virtual reality technology, security experts can more intuitively identify network vulnerabilities, thereby enhancing the evaluation of risks

and threats in the network. Secondly, virtual reality technology enables the simulation of attack scenarios and vulnerability remediation processes [14, 20]. This enhances the skillsets and competencies of security experts. As we progress, we must intensify research into the fusion of virtual reality technology and network security management. This will further enhance virtual reality technology's advantages and guarantee the network environment's stable and secure operation.

**5. Conclusion.** This research introduces the integration of virtual reality technology for computer network security vulnerability detection, addressing issues related to extended detection delays in traditional systems. Through optimization calculations, distinct vulnerability attributes are extracted and analyzed using web crawlers to compile an injection point list. Virtual reality technology simulates attacks, capturing genetic attributes through virtual vulnerabilities. Experimental results exhibit significant reductions in detection time delays, affirming the efficiency of the proposed automatic detection method compared to traditional approaches. The findings reveal a substantial reduction in detection time delays, with respective reductions of 290.11 milliseconds and 337.30 milliseconds compared to the two traditional methods. Virtual reality technology offers substantial application potential in computer network security vulnerability detection. Its ability to simulate real network environments empowers experts to identify and remediate vulnerabilities virtually, enhancing detection precision and efficiency while minimizing risks inherent in real-world assessments. This approach can shape the future of vulnerability detection, accompanying in a new era of effective and efficient security measures within the evolving landscape of network technology.

#### REFERENCES

- [1] A. S. ADEGOKE, T. T. OLADOKUN, T. O. AYODELE, S. E. AGBATO, A. D. JINADU, AND S. O. OLALEYE, *Analysing the criteria for measuring the determinants of virtual reality technology adoption in real estate agency practice in lagos: a dematel method*, Property Management, 40 (2022), pp. 285–301.
- [2] F. BELLALOUNA, *Digitization of industrial engineering processes using the augmented reality technology: Industrial case studies*, Procedia CIRP, 100 (2021), pp. 554–559.
- [3] M. BRUGGER, *Limitations of the hyperplane separation technique for bounding the extension complexity of polytopes*, Operations Research Letters, 49 (2021), pp. 896–901.
- [4] Y. DU, C. DUAN, AND T. WU, *Lubricating oil deterioration modeling and remaining useful life prediction based on hidden semi-markov modeling*, Journal of Engineering Tribology, 236 (2022), pp. 916–923.
- [5] E. FERRETTI, J. R. SCHOENHERR, A. MATTIOLA, AND T. DABOVAL, *Vulnerabilities in clinician–parent exchanges and the cascade of communication traps: a review*, Archives of Disease in Childhood, 108 (2023), pp. 86–90.
- [6] K. P. GUPTA AND P. BHASKAR, *Teachers' intention to adopt virtual reality technology in management education*, International Journal of Learning and Change, 15 (2023), pp. 28–50.
- [7] Y. LI AND X. LI, *Research on multi-target network security assessment with attack graph expert system model*, Scientific Programming, 2021 (2021), pp. 1–11.
- [8] Z. LIN, Y. LAI, T. PAN, W. ZHANG, J. ZHENG, X. GE, AND Y. LIU, *A new method for automatic detection of defects in selective laser melting based on machine vision*, Materials, 14 (2021), p. 4175.
- [9] X. LIU, K. SHI, Z. WANG, AND J. CHEN, *Exploit camera raw data for video super-resolution via hidden markov model inference*, IEEE Transactions on Image Processing, 30 (2021), pp. 2127–2140.
- [10] T.-W. LUI AND L. GOEL, *Learning effectiveness of 3d virtual reality in hospitality training: a situated cognitive perspective*, Journal of Hospitality and Tourism Technology, 13 (2022), pp. 441–460.
- [11] J. MCFADDEN, K. JUNG, B. ROBINSON, AND T. R. TRETTER, *Teacher-developed multi-dimensional science assessments supporting elementary teacher learning about the next generation science standards*, Journal of Science Teacher Education, 33 (2022), pp. 55–82.
- [12] X. MENG, S. KUNDU, A. K. KANUPARTHI, AND K. BASU, *Rtl-contest: Concolic testing on rtl for detecting security vulnerabilities*, IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 41 (2021), pp. 466–477.
- [13] A. O'CONNOR, A. TAI, Z. KOPSAFTIS, AND K. CARSON-CHAHHOUD, *Qualitative study protocol: Augmented reality technology to deliver asthma inhaler technique training for children and adolescents with asthma*, International Journal of Qualitative Methods, 20 (2021), pp. 1–11.
- [14] G. TANG, L. YANG, S. REN, L. MENG, F. YANG, AND H. WANG, *An automatic source code vulnerability detection approach based on kelm*, Security and Communication Networks, 2021 (2021), pp. 1–12.
- [15] Z. WU, C. JIAO, AND L. CHEN, *Tire defect detection method based on improved faster r-cnn*, Journal of Computer Applications, 41 (2021), pp. 1939–1946.
- [16] ———, *Tire defect detection method based on improved faster r-cnn*, Journal of Computer Applications, 41 (2021), pp. 1939–1946.
- [17] X. YANG, L. SHU, J. CHEN, M. A. FERRAG, J. WU, E. NURELLARI, AND K. HUANG, *A survey on smart agriculture: Development modes, technologies, and security and privacy challenges*, IEEE/CAA Journal of Automatica Sinica, 8 (2021), pp. 273–302.

- [18] J. ZHANG, Y. JIN, B. SUN, Y. HAN, AND Y. HONG, *Study on the improvement of the application of complete ensemble empirical mode decomposition with adaptive noise in hydrology based on rbfnn data extension technology.*, CMES-Computer Modeling in Engineering & Sciences, 126 (2021), pp. 755–770.
- [19] D. ZHAO, *Choice of environmental and economic path for building a supply chain financial cloud ecosystem under the background of “internet+”*, Wireless Communications and Mobile Computing, 2021 (2021), pp. 1–11.
- [20] R. ZHENG, H. MA, Q. WANG, J. FU, AND Z. JIANG, *Assessing the security of campus networks: the case of seven universities*, Sensors, 21 (2021), p. 306.

*Edited by:* C. Venkatesan

*Special Issue:* Next Generation Pervasive Reconfigurable Computing for High Performance Real Time Appls

*Received:* Mar 20, 2023

*Accepted:* Aug 30, 2023





## COMPUTER MALICIOUS CODE SIGNAL DETECTION BASED ON BIG DATA TECHNOLOGY

**Abstract.** The article addresses the challenges modelled by the inadequacy of traditional detection methods in effectively handling the substantial volume of software behavior samples, particularly in big data. A novel approach is proposed for leveraging big data technology to detect malicious computer code signals. Additionally, it seeks to attack the issues associated with machine learning-based mobile malware detection, namely the presence of a large number of features, low accuracy in detection, and imbalanced data distribution. To resolve these challenges, this paper presents a multifaceted methodology. First, it introduces a feature selection technique based on mean and variance analysis to eliminate irrelevant features hindering classification accuracy. Next, a comprehensive classification method is implemented, utilizing various feature extraction techniques such as principal component analysis (PCA), Kaehunen-Loeve transform (KLT), and independent component analysis (ICA). These techniques collectively contribute to enhancing the Precision of the detection process. Recognizing the issue of unbalanced data distribution among software samples, the study proposes a multi-level classification integration model grounded in decision trees. In response, the research focuses on enhancing accuracy and mitigating the impact of data imbalance through a combination of feature selection, extraction techniques, and a multi-level classification model. The empirical results highlight the effectiveness of the proposed methodologies, showcasing notable accuracy improvements ranging from 3.36% to 6.41% across different detection methods on the Android platform. The introduced malware detection technology, grounded in source code analysis, demonstrates a promising capacity to identify Android malware effectively.

**Key words:** Android malware detection, Feature extraction, Set classification algorithm, PCA, Kaehunen-Loeve transform (KLT), Independent component analysis (ICA)

**1. Introduction.** With the advent of the first Von Neumann computer and the 20th-century rise of the Internet to today's interconnected society, computers' significance in social life has steadily escalated. Concurrently, as computers have progressed, an unceasing influx of network security incidents has arisen. Such incidents can result in material losses for individuals and organizations and grave consequences, including personal privacy breaches and data leaks. According to the security report spanning August 2021 to January 2022 by the National Internet Emergency Response Center, a statistical chart depicting security events becomes evident. Over this half-year period, malicious program events constituted nearly a third of China's total security incidents, consistently ranking among the top two alongside vulnerability events in network security [5]. In contemporary society, the Internet bears substantial social and economic value. Yet, specific unlawful accesses enable illegal actors to hijack computers, private data, etc., yielding economic or political gains. A case in point is the WannaCry ransomware, which encrypts and extorts data from hospitals, banks, and other institutions, and the Stuxnet worm, which increased in Iran to target critical energy facilities. The former encrypts the victim host computer data and forces users into paying for data decryption. At the same time, the latter disrupts uranium plant centrifuges through a computer virus, indirectly impacting Iran's nuclear program to achieve political objectives [2].

Traditional malware detection approaches centered around feature matching and heuristic scanning fall short in adapting to the dynamic evolution of malware. Modern malicious software frequently employs tactics like shell techniques and encryption to elude conventional detection methods, necessitating more advanced technologies for effective detection. The robust capacity of machine learning to discern latent correlations between features has prompted an increasing number of experts to integrate machine learning techniques into malware identification. The malware detection approach rooted in machine learning capitalizes on conventional static and dynamic detection mechanisms to extract malware attributes. These attributes are subsequently translated into formats like images and text, allowing for identification through classical machine learning classification algorithms, convolutional neural networks (CNNs), recurrent neural networks (RNNs), and other methodologies, yielding commendable outcomes. Recent years have witnessed notable strides in entity relationship extraction,

primarily due to the advancement of graph convolutional networks (GCNs). Scholars have embarked on incorporating graph neural networks into pertinent domains, notably including the domain of malware detection. Notably, control flow graphs (CFGs) and function call graphs (FCGs) of binary files have emerged as pivotal starting points in this research direction within the malware detection field [10, 11].

The paper is structured into five main sections. Section 1 introduces the study's focus on computer malicious code signal detection using big data technology. Section 2 presents a comprehensive literature review to contextualize the research. In Section 3, the proposed method for detection is detailed. Section 4 encompasses the presentation and discussion of the results obtained. Finally, Section 5 provides the concluding remarks summarizing the findings and contributions of the study.

**2. Literature Review.** A load decomposition system is introduced that efficiently connects high-frequency data from battery-powered particles equipped with current sensors, complementing the infrastructure's low-frequency data acquired from energy meters [12]. Machine learning techniques are employed to classify executable files, including a k-nearest neighbor, support vector machine (SVM), and decision tree methods. The effectiveness of this approach was measured against dynamic detection methods that utilize dynamic system call sequences. Empirical findings indicate that this approach demonstrates enhanced detection accuracy and a reduced false positive rate compared to dynamic detection methods [3]. A malware detection approach is discussed in risk theory, incorporating feature extraction and synthesis. Using the n-gram algorithm, features are extracted from API call sequences during malware runtime and consolidated into risk and security signals. Ultimately, the deterministic dendritic cell algorithm is employed for malware detection. Experimental results highlight that, in contrast to four other detection algorithms, the proposed method displays a lower false positive rate and false negative rate [7].

Since Kaspersky discovered SMS Trojans in August 2010, Android malware has evolved rapidly. Delving into Android malware reveals that the latest malware families exhibit diverse malicious behaviors, encompassing encrypted payloads, code obfuscation, concealed commands, and manipulation of communication channels. As cutting-edge technologies advance and the black market's allure grows, the design of Android malware has become increasingly intricate and ingenious. To safeguard user information, a surge of technologies and research efforts are being channeled into malware prevention and detection [14].

Regarding feature extraction methods, two primary approaches prevail: static analysis and dynamic analysis. The dynamic detection approach involves executing Android software within controlled environments like virtual machines or sandboxes, observing software behavior and entire execution processes to discern potential malicious activities. This approach is capable of identifying anomalies in zero-day vulnerabilities. However, it might fail to induce malicious behavior in spyware adept at evasion. Additionally, dynamic detection necessitates substantial hardware resources, system allocations, and extended detection cycles.

Conversely, static detection for Android malware typically entails scrutinizing software source code or decompiled executable files to extract features and achieve malware detection without code execution. Static detection is a software analysis technique that uncovers defects or vulnerabilities by analyzing source code or executable files. In contrast to dynamic detection, static detection doesn't demand code execution, allowing it to identify potential issues and rendering it more practical rapidly. This method is widely adopted by researchers in the Android malware detection domain due to its speed, efficiency, and minimal resource requirements [15].

Many specialized big data-driven techniques for detecting malware are designed explicitly for mobile software operations. This encompasses a comprehensive exploration of Dalvik, API, and permission-based profiling of malware targeting Android systems. The investigation explores deeply into feature selection, covering frequency-averaged and variance-based algorithms. Moreover, the study extensively explores feature extraction methods, including PCA, KLT, and ICA. The research also examines a decision tree-powered multi-level classification fusion algorithm strategically devised to rapidly and accurately identify instances of mobile malware within the Android platform [18].

**3. Proposed Methodology.** Mobile malware detection systems rely on specialized selection algorithms rooted in mean and variance analysis, content extraction-driven malware detection algorithms, and multi-level mixture-based malware detection algorithms. A mobile malware detection system serves the purpose of identifying malicious software on mobile devices. Given the escalating integration of mobile devices into

everyday activities and the escalating risk posed by malicious software, developing such a detection system holds immense importance [9, 13]. The function of each component within the system is delineated as follows.

- 1) Feature extraction: It extracts the Permission feature, API feature and Dalvik feature from the collected samples.
- 2) Set learning: multi-level integration is adopted for classification and fusion [22, 17].

**3.1. Mean-Variance-based feature selection algorithm.** The Dalvik directive is in the Android software's runtime on the Dalvik virtual machine, offering information that lies deeper within the system than the API. It presents the Android software's implementation through its registry, offering the benefits of reduced instructions and simplified withdrawal, rendering it adept for effective malware detection. This section explores the efficacy of Dalvik's specialized instructions for detecting malware. During the compilation of an Android app, Java bytecode is automatically transformed into instructions tailored for the Dalvik VM, stored within the class dex. This classes.dex file is intrinsic to every Android application and is executable directly on the Dalvik VM. During runtime, the Dalvik VM accesses register operands using Dalvik instructions, encompassing null operation (nop), data operation (move), return, data definition (const), method call (invoke), data conversion, and data operation instructions. This paper's approach relies on representing the relative frequency of Dalvik instructions and employs a feature selection algorithm [1].

Disparities in the relative frequencies of Dalvik instructions hold significant implications for software vulnerability assessment. The mean relative frequency of specific Dalvik instructions within malware surpasses the normal software, indicating that malware frequently invokes specific Dalvik instructions more often than regular software. Moreover, the instruction divergence among certain malware instances surpasses that within standard software, underscoring distinct calling patterns for Dalvik instructions among different malware strains. This paper introduces Dalvik message features, involving averaging and numeric system selection.

The extent of divergence between malware and normal software regarding the relative frequency of invoking Dalvik instructions correlates with enhanced performance. Simultaneously, improved proximity between sample points of the same category signifies better relative frequency within samples of the same classification, alongside greater separation between sample points of disparate categories. The paper incorporates a linear discriminant analysis (LDA) formula to assess the classification potency of diverse Dalvik instructions, as illustrated in Equation (3.1).

$$J(A) = \frac{|\mu_b - \mu_m|^2}{S_b^2 + S_m^2} \quad (3.1)$$

where  $\mu_b$  and  $\mu_m$  represents the average value of the relative frequency of standard software and malicious software, respectively;  $S_b$  and  $S_m$  represents the variance of standard and malicious software's relative frequency. It can be seen from Equation (3.1) that the goal of feature selection is to maximize the  $J(A)$  value, that is, to minimize the sum of variances between the two types of samples and to maximize the difference between the average values of the two types of samples. We calculate the  $J(A)$  value for each Dalvik instruction. The larger  $J(A)$  value is, the stronger the classification ability of feature  $A$ , and sort the features in descending order according to the size of the  $J(A)$  value. Finally, the first  $k$  Dalvik instruction features is selected as a practical feature subset [21, 6].

**3.2. Feature extraction-based malware detection algorithm.** The malware detection algorithm centered around feature extraction employs machine learning and other techniques to identify malware through feature analysis. The fundamental concept involves initially extracting various attributes from malicious software, such as API call sequences, permission utilization, and code structure. Subsequently, these features are fed into a classifier for training, yielding a model capable of recognizing malicious software. Eventually, this model is utilized to categorize unfamiliar software. A content extraction algorithm highlights distinctions in features across diverse locations to enhance detection accuracy. Figure 3.1 illustrates the proposed novel malware detection strategy based on feature extraction.

Initially, the extraction of Dalvik instructions from the code is performed. Subsequently, employing fundamental analysis, KLT, and ICA, the original Dalvik instructions are mapped to corresponding positions, generating three novel instructions. Lastly, a single-layer neural network is trained using a rapid machine

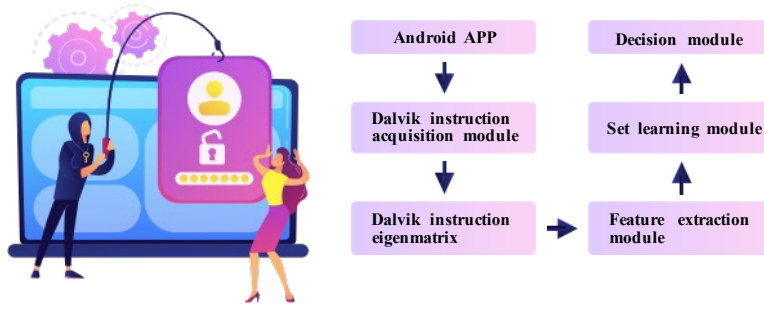


Fig. 3.1: Flow diagram of malware detection system based on feature extraction

learning algorithm as the foundational classifier. The stacking process then integrates the overall distribution base [4].

The three distinct feature extraction algorithms used to map features into distinct feature spaces are as follows:

(1) PCA: It is a linear projection technique that maps high to low dimensional features, and it is a widely used data reduction technology. Its purpose is to extract the most essential features from high-dimensional data and convert them into vectors in low-dimensional space to realize data visualization and simplify analysis. In the process of projection, the variance of data is the largest. Define  $N$  samples,  $X = [X_1, X_2, \dots, X_N]$  where;  $X_i = [X_{i,1}, X_{i,2}, \dots, X_{i,d}]^T \in \mathbb{R}^d$ ;  $d$  is the number of features. The principal component analysis transforms linearly into a new sample, as follows in Equation (3.2):

$$y_i = U^T X_i \quad (3.2)$$

where  $U$  is an orthogonal matrix. PCA initially computes the sample's covariance matrix. Subsequently, it determines the eigenvectors of this covariance matrix and transforms the original features.

(2) Kaehunen-Loeve transform: In PCA, the transformation matrix is the covariance matrix of the sample. In the Kaehunen-Loeve transformation, the transformation matrix is the inter-class dispersion matrix, which is recorded as  $S_w$  given in Equation (3.3):

$$S_w = \sum_{i=1}^L P_i E \left[ (X - \bar{m}_i) (X - \bar{m}_i)^T \right] \quad (3.3)$$

where  $L$  is the sample category;  $P_i$  is the probability of category  $i$ ;  $E$  stands for mathematical expectation;  $m_i$  is the mean of category  $i$ . Calculate the eigenvector of  $S_w$ , and then calculate the new feature according to formula (3.2).

(3) Independent component analysis: The independent information is extracted from original features. The independent component analysis model is recorded using Equation (3.4):

$$X = As \quad (3.4)$$

where  $X$  is the original data;  $A$  is a full rank matrix,  $S$  is an independent component. The purpose of independent component analysis is to extract independent component  $S$  from  $X$  is represented in Equation (3.5):

$$\hat{s} = UX \quad (3.5)$$

where  $\hat{s}$  represents the estimated value of the independent component;  $U$  is the transformation matrix and fast ICA algorithm is used to calculate  $U$ .

**3.2.1. Set classifier based on extreme learning machine.** Utilizing the extreme learning machine (ELM) algorithm, the single-layer neural network is employed as a foundational classifier. The Stacking technique is then employed to construct a set classifier. Noteworthy advantages of the ELM ensemble classifier encompass:

1. Swift Training Pace: The ELM algorithm's training process entails solving a linear equation system and does not require iterative optimization. Consequently, it achieves rapid training, proving especially suited for extensive datasets.
2. Elevated Classification Precision: The ELM algorithm boasts robust generalization capabilities and is adept at addressing intricate scenarios like high dimensionality, nonlinearity, and noise interference, resulting in high classification accuracy.
3. Robust Scalability: The ELM algorithm integrates with other machine learning methods, such as SVM and random forests. Its applicability extends to multi-sample, multi-feature, and multi-modal datasets.

The swift learning machine algorithm randomly assigns initial input weight vectors and biases to the neural network. Subsequently, the neural network's output weights are determined through analysis. The algorithmic sequence is as follows:

In the training stage, input weights and deviations are randomly assigned to calculate the output of hidden layer nodes, as shown in the following formula (3.6):

$$h_{ij} = g(w_j x_i + b_j) \quad i = 1, 2, \dots, N; j = 1, 2, \dots, k \quad (3.6)$$

where  $h_{ij}$  is the output of the  $j$  hidden layer node;  $w_j = [w_{j1}, w_{j2}, \dots, w_{jn}]^T$  is the weight value connecting the  $j$ -th hidden node and the input data;  $b_j$  is the node deviation of the  $j$ -th hidden layer;  $N$  is the number of samples;  $k$  is the number of hidden layer nodes;  $g$  is the activation function. The output matrix of the hidden layer is marked as  $H = \{h_{ij}\}$ , and the weight vector connecting the hidden layer and the output layer nodes is marked as  $\beta$ , as follows (3.7):

$$\hat{\beta} = H^\dagger T \quad (3.7)$$

where  $\hat{\beta}$  is the estimated value of  $\beta$ ;  $H^\dagger$  is the generalized inverse matrix of  $H$ 's Moore-Penrose;  $T$  is the classification label as given by Equation (3.8). In the test phase, for unknown samples, the hidden layer node output  $H$  is calculated first, and then its label is predicted.

$$T = \hat{\beta} H \quad (3.8)$$

where  $T$  is the prediction label of unknown samples;  $H$  is the hidden layer output of unknown samples, calculated according to Equation (3.6). This paper uses the Stacking method to build a set classifier and stacking fuses of each base classifier based on the learning method, including two layers marked as level-0 and level-1. In level-0, an extreme learning machine builds  $N$  basis classifiers. In level-1, the tags predicted in level-0 are used as input data, and the speed learning machine is also used for training [16].

**3.3. Multi-level integration-based algorithm for mobile malware detection.** Android software permissions and API functions are pivotal in multi-level integrated mobile malware detection. An in-depth analysis uncovers distinctions in the frequency of usage of permitted functions and API functions between normal software and malware. Hence, this article selects and compares the top 20 permissions and APIs, with the outcomes presented in Figures 3.2 and 3.3. In practical scenarios, the quantity of genuine software far exceeds that of malicious software, creating an inequality in data distribution. Generally, approaches addressing uncertain data involve resampling, large-sample size reduction, SVM, and Random Forest. Primarily, the focus centers on detecting malware, endeavoring to minimize instances of malware being wrongly identified as normal software. This involves diminishing false positives. Consequently, the prevailing approach involves customizing diverse training and combination methodologies to cater to distinct models [19].

It can be seen from Figure 3.2, the normal software and malicious software have a high frequency of applying for permissions, such as WRITE\_EXTERNAL\_STORAGE and ACCESS\_NETWORK\_STATE, but the difference between them is small; However, the frequency of malicious software applications for

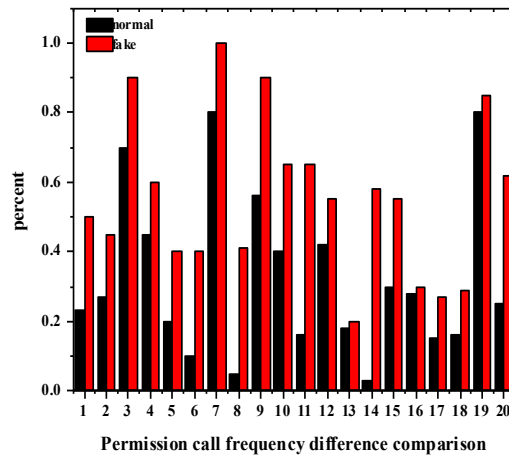


Fig. 3.2: Comparing the disparity in call frequency between normal and malicious software

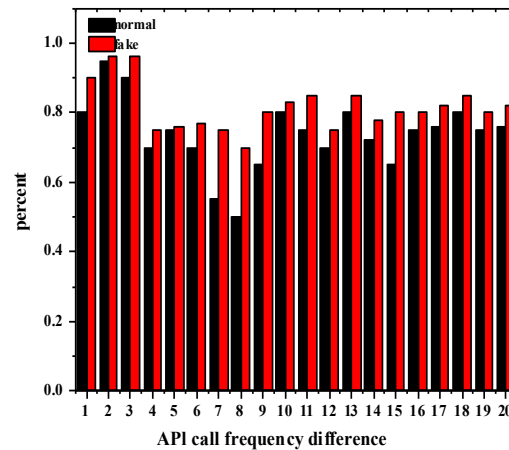


Fig. 3.3: Contrasting API call frequencies in normal and malicious software

READ\_CONTACTS, SEND\_SMS and other permissions is much higher than that of normal software applications. Therefore, permission is effective as a feature to distinguish normal software from malicious software.

The observation from Figure 3.3 indicates that both normal and malicious software frequently raise APIs like `Android.content.Content` and `android.app.Activity`. Despite this, the disparity between their usage frequencies remains minimal. In contrast, malware exhibits a higher frequency of invoking APIs such as `android.graphics.Paint` and `android.View.KeyEvent` compared to the frequency seen in normal software [8].

To address the issues as mentioned earlier and enhance detection accuracy, we present a novel approach known as decision tree ensembles-based detection (DTED) method. This integration technique involves a three-tier ensemble of decision trees. Among the decision tree-based integrated learning methods, the random forest algorithm is renowned. Initially, the random forest algorithm extracts a specific portion of samples and features randomly from the original data, constructing multiple decision trees, each trained on distinct sample and feature sets. During classification, test samples are input into each decision tree for classification, with the outcomes from each decision tree being collectively voted upon or subjected to weighted averaging to derive the ultimate classification result. In regression scenarios, the prediction outcomes of individual decision trees are amalgamated using an averaging approach. The framework for mobile malware detection is proposed using

Table 4.1: Composition of datasets

Dataset	Malware	Regular software	Total
D1	29102	29120	58222
D2	1262	1262	2524

the following three components:

- (1) The initial set of combinations employs the complete voting approach, ensuring an equitable distribution across all decision trees. This balance mitigates ambiguous information and addresses the issue at hand.
- (2) The subsequent stage of amalgamation employs a majority voting model. For models not identified in the first stage, they are combined with the majority vote of the first-stage decision trees. This integration occurs alongside the processing of different effective models from the second stage. This technology effectively reduces malware false alarms.
- (3) The third layer of integration leverages a select few votes. Typically, following the two preceding detection layers, the remaining undetected samples tend to be more resilient. In such cases, where robust feature characterization is lacking, incorporating a few experienced votes can enhance accuracy to a certain extent.

Compared with classical data mining algorithms, DTED effectively addresses unbalanced data, offers high accuracy, and maintains a low false alarm rate. This success stems from the fact that the first-level integration method can effectively detect most positive samples, resulting in a balanced proportion of positive and negative samples, thus tackling the imbalance challenge. The second-level integration, combined with the first-level integration, bolsters the weight of negative samples, enhancing their detection rate and thereby diminishing the false alarm rate of negative samples. The third integration layer fortifies the outcomes from the second tier, heightening the overall detection accuracy. In summary, the DTED method excels in handling imbalanced data, and delivering high accuracy.

#### 4. Results and Discussion.

**4.1. Data sets and evaluation indicators.** The D1 dataset encompasses models from Google Play, Peapod, and Android Online, totaling 29,102 models. On the other hand, the malware samples originate from the Android Malware Genome Project and Andro MalShare, amounting to a total of 29,120 samples. Malware samples for the D2 installation file dataset are sourced from the Android Malware Genome Project, while the source software exclusively comes from Google Play. Notably, dataset D1 is utilized for testing the feature selection algorithm, whereas dataset D2 serves as the testing ground for the feature extraction algorithm, as illustrated in Table 4.1.

Numerous classification systems exhibit distinct characteristics, with diverse attributes across different methods classes and performance outcomes varying based on data types. The realm of classification algorithms encompasses an array, including decision trees, naive Bayes, SVM, logistic regression, random forest, neural networks, and more. Each algorithm possesses unique merits, drawbacks, and suitable contexts, with algorithm selection contingent upon specific circumstances. The choice of classification hinges on the task's nature, the methodology for classification selection, and the approach to evaluating algorithm efficacy. Generally, Precision and F-measure stand as chosen measurement parameters. To begin, we will introduce various evaluation models. For now, let's consider our target categories as positive and negative groups [20].

1. True positive (TP): The number of events that are positive events, divided by classification into good conditions;
2. False positive (FP): the number of cases misclassified by positive cases, i.e. divided by the number of cases classified as negative but in the excellent condition;
3. False negative (FN): the number of cases classified as negative, i.e., the number of cases that were good but classified as negative by category;

Table 4.2: Classification accuracy of different feature selection algorithms

Classifier	Accuracy (%)				
	No feature selection	Sequential Forward Selection	Mountain climbing algorithm	Relief algorithm	Proposed Feature Selection based on Mean and Variance
Logistic regression	81.57	63.57	78.81	82.43	82.70
C4.5 Decision tree	87.04	77.75	86.95	87.53	87.61
Random forest	83.98	79.81	82.26	90.39	90.56

Table 4.3: Comparative performance of classifiers based on accuracy (%) and F-measure (%)

Classifier	Accuracy (%)	F-measure (%)
ELM	93.56	93.43
ELM+PCA	92.10	93.81
ELM+KLT	93.78	90.73
ELM+ICA	93.55	92.95
Stacking	95.12	93.58
Proposed FES	97.52	97.52

4. True negative (TN): The number of events correctly classified as negative events, that is, the number of negative events divided by negative events by category. The classification criteria are as follows.
- Precision: Precision is a measure of accuracy, which represents the proportion of positive cases in the cases divided into positive cases, "Precision"=TP/(TP+FP);
  - Recall rate: recall rate is the degree of coverage, "Recall"=TP/(TP+FN);
  - F-measure: F-measure is the average harmonic number of recall and precision is expressed as Equation (4.1):

$$\text{F-measure} = \frac{2 \times \text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} \quad (4.1)$$

**4.2. Performance comparison of various classifiers.** To evaluate the proposed FSMV classification system, this study employs regression logic, C4.5 decision trees, and random forest as classifiers, comparing them with subsequent selection algorithms (SFS), hill climbing, and Relief algorithms. Therefore, the classification outcomes using D1 data are presented in Table 4.2. The data in Table 4.2 demonstrates the enhancement in the detection capabilities of various classification systems facilitated by the specialized selection algorithm based on the mean and variance (FSMV) of Dalvik instructions. The combination of FSMV with random forest for detecting Dalvik definitions showcases a 6.41-fold improvement over the sole use of the random forest algorithm. This improvement lies below the Relief algorithm's performance and surpasses that of the C4.5 decision tree when identifying regression logic connections accurately.

To test the malware system based on the extracted content, this paper uses the D2 dataset, randomly selecting 80 files as training data and 20 files as testing data. The corresponding results are shown in Table 4.3. The table shows that the F-value of ELMPCA is higher than that of the ELM algorithm alone, and the F-value of ELMKLT and ELMICA is lower than that of ELM alone; The accuracy and F-rate of using only Stacking is higher than that of using ELM only, and the content extraction-based Stacking method is higher than other methods, indicating that it performs well.

To evaluate the malware detection method based on multi-level integration, this study employs dataset D1, randomly selecting 80% of the data for training and 20% for testing. Typically, in imbalanced data, low



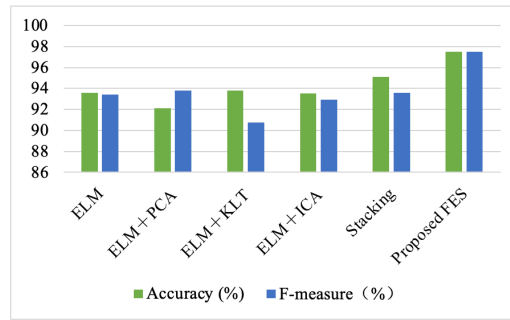


Fig. 4.1: Comparison of classification accuracy and F-measure for different classifiers and feature extraction techniques

Table 4.4: Classification accuracy and F-measure value of different algorithms

Detection algorithms	Accuracy (%)	F-measure (%)
J48 decision tree	87.37	88.24
SVM	88.13	87.20
Random forest	89.87	89.15
Proposed DTED	91.60	91.13

omission rate, and accuracy requirements, the F-measure value is widely employed as the evaluation metric. The results are displayed in Table 4.4.

The provided table presents the outcomes of classification accuracy and F-measure achieved through diverse classifiers and feature extraction techniques. The ELM classifier achieves an accuracy of 93.56% and an F-measure of 93.43%. When coupled with different feature extraction methods, ELM yields variable outcomes. “ELM+PCA” demonstrates an accuracy of 92.10% and an F-measure of 93.81%, “ELM+KLT” achieves an accuracy of 93.78% but a comparatively lower F-measure of 90.73%, and “ELM+ICA” results in an accuracy of 93.55% and an F-measure of 92.95%.

Remarkably, the “Stacking” technique achieves an accuracy of 95.12% and an F-measure of 93.58%. Nevertheless, it’s the feature extraction system (FES) that boasts the highest accuracy at 97.52%, coupled with an impressive F-measure of 97.52%. These outcomes underscore the effectiveness of distinct classifiers and their combinations with varied feature extraction techniques in terms of accuracy and F-measure for malware detection. Among these, the “Proposed FES” strategy emerges as the most accurate classifier in this particular context.

The table shows that the proposed DTED method surpasses commonly utilized solutions for addressing imbalanced data, such as SVM and random forest. Notably, the DTED method exhibits higher accuracy and F-measure values than alternative approaches. This outcome highlights that the DTED method excels in handling uneven data and meeting the demands for low omission rates and high accuracy. Thus, the DTED method showcases a strong and effective detection performance.

**5. Conclusion.** The rapid expansion of the mobile Internet has facilitated the explosion of mobile malware, mainly targeting the Android platform due to its open nature. Consequently, evaluating the security of applications released on the internet and app stores and discriminating malicious software remains a pressing research concern. This study critically examines the current challenges of mobile terminal malware detection technology. It introduces a novel approach by scrutinizing software attributes from a source code perspective while investigating pivotal facets of Android malware detection, encompassing Dalvik instructions, permissions, and APIs. Essential techniques such as feature selection and extraction algorithms, alongside multi-level

integrated classification methods, were meticulously explored. Empirical findings validate the proposed source-based malware detection technology, confirming its efficiency in accurately identifying Android malware. This research contributes substantially to the ongoing discourse by presenting a multi-layered solution that addresses the growing threat of mobile malware. Connecting source-based detection and comprehensive methodologies offers a robust safeguard for the Android platform's security and integrity in the evolving mobile Internet.

## REFERENCES

- [1] M. M. ALANI AND A. I. AWAD, *Paired: An explainable lightweight android malware detection system*, IEEE Access, 10 (2022), pp. 73214–73228.
- [2] S. M. BELLOVIN, S. LANDAU, AND H. S. LIN, *Limiting the undesired impact of cyber weapons: technical requirements and policy implications*, Journal of Cybersecurity, 3 (2017), pp. 59–68.
- [3] Y. DING, X. YUAN, D. ZHOU, L. DONG, AND Z. AN, *Feature representation and selection in malicious code detection methods based on static system calls*, Computers & Security, 30 (2011), pp. 514–524.
- [4] E.-S. M. EL-KENAWY, S. MIRJALILI, F. ALASSERY, Y.-D. ZHANG, M. M. EID, S. Y. EL-MASHAD, B. A. ALOYAYDI, A. IBRAHIM, AND A. A. ABDELHAMID, *Novel meta-heuristic algorithm for feature selection, unconstrained functions and engineering problems*, IEEE Access, 10 (2022), pp. 40536–40555.
- [5] D. A. FERNANDES, L. F. SOARES, J. V. GOMES, M. M. FREIRE, AND P. R. INÁCIO, *Security issues in cloud environments: a survey*, International Journal of Information Security, 13 (2014), pp. 113–170.
- [6] A. HASHEMI, M. B. DOWLATSHAHI, AND H. NEZAMABADI-POUR, *Ensemble of feature selection algorithms: a multi-criteria decision-making approach*, International Journal of Machine Learning and Cybernetics, 13 (2022), pp. 49–69.
- [7] C. HUANG, J. CHEN, S. GONG, Q. LUO, AND Q. ZHU, *Feature representation and selection in malicious code detection methods based on static system calls*, Journal of Central South University(Science and Technology), 45 (2014), pp. 3055–3060.
- [8] U. M. KHAIRE AND R. DHANALAKSHMI, *Stability of feature selection algorithm: A review*, Journal of King Saud University-Computer and Information Sciences, 34 (2022), pp. 1060–1073.
- [9] J. KIM, Y. BAN, E. KO, H. CHO, AND J. H. YI, *Mapas: a practical deep learning-based android malware detection system*, International Journal of Information Security, 21 (2022), pp. 725–738.
- [10] T. KIM, B. KANG, M. RHO, S. SEZER, AND E. G. IM, *A multimodal deep learning method for android malware detection using various features*, IEEE Transactions on Information Forensics and Security, 14 (2018), pp. 773–788.
- [11] V. KOULIARIDIS AND G. KAMBOURAKIS, *A comprehensive survey on machine learning techniques for android malware detection*, Information, 12 (2021), p. 185.
- [12] M. KUMAR, *Scalable malware detection system using big data and distributed machine learning approach*, Soft Computing, 26 (2022), pp. 3987–4003.
- [13] S. R. T. MAT, M. F. AB RAZAK, M. N. M. KAHAR, J. M. ARIF, AND A. FIRDAUS, *A bayesian probability model for android malware detection*, ICT Express, 8 (2022), pp. 424–431.
- [14] M. MIJWIL, I. E. SALEM, AND M. M. ISMAEEL, *The significance of machine learning and deep learning techniques in cybersecurity: A comprehensive review*, Iraqi Journal For Computer Science and Mathematics, 4 (2023), pp. 87–101.
- [15] A. MITRA, B. BERA, A. K. DAS, S. S. JAMAL, AND I. YOU, *Impact on blockchain-based ai/ml-enabled big data analytics for cognitive internet of things environment*, Computer Communications, 197 (2023), pp. 173–185.
- [16] W. QIAN, Y. XIONG, J. YANG, AND W. SHU, *Feature selection for label distribution learning via feature similarity and label correlation*, Information Sciences, 582 (2022), pp. 38–59.
- [17] J. QIU, Q.-L. HAN, W. LUO, L. PAN, S. NEPAL, J. ZHANG, AND Y. XIANG, *Cyber code intelligence for android malware detection*, IEEE Transactions on Cybernetics, 53 (2022), pp. 617–627.
- [18] S. SAHOO, *Big data analytics in manufacturing: a bibliometric analysis of research in the field of business management*, International Journal of Production Research, 60 (2022), pp. 6793–6821.
- [19] R. SUDHARSAN AND E. GANESH, *A swish rnn based customer churn prediction for the telecom industry with a novel feature selection strategy*, Connection Science, 34 (2022), pp. 1855–1876.
- [20] A. THAKKAR AND R. LOHIYA, *A survey on intrusion detection system: feature selection, model, performance measures, application perspective, challenges, and future research directions*, Artificial Intelligence Review, 55 (2022), pp. 453–563.
- [21] Y. WU, J. SHI, P. WANG, D. ZENG, AND C. SUN, *Deepcatra: Learning flow-and graph-based behaviours for android malware detection*, IET Information Security, 17 (2023), pp. 118–130.
- [22] P. YADAV, N. MENON, V. RAVI, S. VISHVANATHAN, AND T. D. PHAM, *A two-stage deep learning framework for image-based android malware detection and variant classification*, Computational Intelligence, 38 (2022), pp. 1748–1771.

*Edited by:* C. Venkatesan

*Special Issue:* Next Generation Pervasive Reconfigurable Computing for High Performance Real Time Apps

*Received:* Mar 20, 2023

*Accepted:* Sep 1, 2023



## BIG DATA IN HEALTHCARE – A COMPREHENSIVE BIBLIOMETRIC ANALYSIS OF CURRENT RESEARCH TRENDS

**Abstract.** The primary purpose of this study is to perform a comprehensive bibliometric analysis of research landscape of big data in healthcare. Big data as a significant technology used in healthcare during the past decade has led to the exponential growth in scientific literature. This study is focused on analysis of many crucial bibliometric indicators such as, overall research output, author productivity, institutional productivity, country wise productivity, collaboration analysis, research trends along with a thematic focus of research output in big data and healthcare. The analysis has been performed on 2294 research articles published in 1018 publication sources from SCOPUS and Web of Science databases. The initial results of the study performed from year 2012 reveals that in the first year 6 research articles were published in the given domain. Then every year the growth of published articles in the field was exponential, however years 2019 to 2021 were the most productive and incremental in terms of number of publications. The analysis results of the study present the performance analysis of research production in terms of annual scientific production, most globally cited articles, author's production over the time, most productive countries, and most relevant affiliations. In addition, the science mapping analysis including the indicators such as, keyword Co-occurrence, Thematic Mapping, Most Relevant Authors, annual source distribution, and collaboration Network analysis has been presented. The study delivers expedient contribution to the field of study by noticeably offering comprehensive analysis results regarding research hotspots and trends, thematic emphasis, and future direction of research in the field. These outcomes will aid researchers in big data and healthcare in planning and designing the research and the challenges and opportunities needed to be explored.

**Key words:** bibliometric analysis, big data, healthcare, thematic mapping, KCN, Collaboration networks

**1. Introduction.** Knowledge not only aids in better understanding a subject or a situation, but it also helps in making better decisions. The more data leads to the better decisions and the better results. In the present technological era where almost everything is digitized [1], a large number of applications create a huge amount of data. International Data Corporation (IDC) estimates that the global data sphere will reach 175 Zettabytes by 2025, in its Data Age 2025 research for Seagate[2]. This situation pushed technocrats and large companies to design the solutions which not only manage the data but also retrieve meaningful information out of that data. Eventually these solutions should fulfill both current and future demands irrespective of the domains. There is no doubt that healthcare organizations have major contribution in data generation, making the industry extremely data-intensive, and electronic data has played an increasingly important role in understanding and improving health [3]. Like other application domains, healthcare too has varieties of data which makes big data in healthcare intimidating and the velocity of healthcare data demands effective and efficient data management solutions. Big data scientists are in high demand because they can assist in uncovering relationships and analyzing patterns and trends in data. Big data analytics has the potential to enhance care, save lives, and reduce costs. As a result, big data analytics apps in healthcare take advantage of the massive amount of data available to extract insights and make better decisions [4]. Big data will assist the healthcare industry in developing more comprehensive and insightful diagnosis and treatments, resulting in greater quality care at lower costs and better overall outcomes.

Thus the enormous need for big data has prompted researchers, academics, and professionals to delve deeper into it, as seen by the large number of research studies performed in the near past at a quick pace. It is getting increasingly difficult to keep up with the updated body of knowledge in the field of study as the research outputs are published at an exponential rate [5]. The objective of this study is to combine existing references with bibliometric analysis to show the status and development of research in big data in healthcare, as bibliometrics is one of the best approaches that use statistical tools to assess research articles and other types of publications over time.

To achieve the primary goals of this study performance analysis and scientific mapping analysis have been performed in this study. Many important bibliometric indicators have been analyzed such as, scientific article

production, total citations per year, most productive authors, most internationally cited documents, most cited nations, most relevant affiliations, most relevant keywords, and most relevant authors are all included in the performance analysis. In addition science mapping analysis has been carried out to analyze the bibliometric indicators such as, keyword co-occurrence, coupling map, theme mapping, and author, institution and country collaboration networks.

#### Highlights:

- To provide a comprehensive bibliometric analysis reference workflow for an interdisciplinary field of study.
- To provide the analysis of evolution and research trends of a field of study such as big data and healthcare.
- To analyze the research productivity in terms of countries, Institutions and authors in the field.
- To analyze the collaborative relationships among countries, institutions and authors in the research field.
- To provide science mapping and thematic analysis for big data and healthcare domain.

**2. Related Work.** In the present technological era every hour, a massive amount of data is generated from almost every environment, big or small. The data is considered a crucial asset of every organization and field. In other words, the more detailed data we have, the more optimally the operations and decisions can be made by the organization, Aria et al. [5]. Considering such critical importance of data, we are flooded by massive amounts of data from every organization. Information and communication technology (ICT) advances have significantly contributed towards generating and collecting the amounts of data to an extreme where it became very challenging to manage it. To deal with the massive amounts of data related to research studies, their results, contributions, and impact, technologies such as big data analysis and Text mining have been effectively used in recent years. Text mining is a method of analyzing a collection of documents as a knowledge-intensive task to recognize and discover interesting patterns Guandong Son et al. [6]. Bibliometrics is a quantitative information assessment method to analyze the emerging developments in a desired field of study to find measurable research output.

Guo, Yuqi, et al. in [7] have proposed a bibliometric search strategy to provide critical insights in the research related to the application of Artificial Intelligence in health care. The study has used the databases of Web of Science, such as the Science Citation Index (SCI) and the Social Science Citation (SSCI). The authors have conducted the temporal and spatial bibliometric analysis. In addition, the analysis has been done based on word co-occurrence, co-country, and co-authorship.

Dash, S et al. [8] provide a comprehensive review of big data management and analysis in healthcare. The results of the research review have proven the potential of big data application in better clinical prognosis. In addition, the application of big data analysis has been proven as an effective tool in reducing cost in analytics, making clinical decision support systems, making superior strategies of treatment as well as detecting and avoiding the fraudulent associated with data.

Galetsis, P et al. [9] present the applicability of big data Analytics in healthcare in terms of nature and scale of innovations in information processing and analysis tools. The study also explored the impact of technological tools and the possible information sources. To achieve the goals, bibliometric analysis has been done on the data sources extracted from Web of Science and Scopus databases. The findings of the study report a massive amount of work done published in numerous research papers related to oncology and neurology. The study thus reports the possible usability of big data analytics in advanced health information and Decision support systems. The findings further prove that the analysis tools will provide the solutions for sophisticated disease prognosis and diagnosis systems.

The application of data mining in healthcare and more specifically in medicine has been extensively researched in recent times throughout the world Hu, Yuanzhang et al. [10]. There is a significant need for research to provide a clear picture of growth and development trends of data mining and analytics in the field using bibliometrics. The study has proposed a bibliometric analysis example to give a clear overview of applying data mining methods in medicine. Various visualization tools have been used to analyze the citations, research collaborations, and spatial dissemination related to data mining in medicine.

Ale Ebrahim S et al. [11] have used a bibliographic method to study the research trend in drug delivery. The

data has been sourced from SCOPUS to examine the research trend for almost 45 years till 2019. The network analysis method has been used by the study for research output analysis. The bibliographic analysis has been done on journal research articles in terms of citations, country-wise contribution based on specific keywords and topics. The bibliometric reports have been used to study the present status and required upcoming research directions in drug delivery.

Wu Haiyang et al. [12] have performed a study aimed at the identification of research output and impact of ultrasound micro bubble as a therapeutic method for diseases such as cancer and neurological and cardiovascular disorders. The literature related to the field of study has been used for 20 years till 2019 extracted from sources like Web of Science Core collection and also SCOPUS data has been used for invalidation. The bibliometric results such as a number of publications, document citations, journal citations, H-Index, authorship, co-authorship, country wise, institution wise, and keyword-based analysis have been performed using various data processing and visualization tools. The study has reported an increased amount of research results supporting the applicability of the ultrasound bubble in worldwide trends generally, however, the United States is leading in the field. The trends in the study depicted the significant research needs for the possibility of ultrasound usability in drug delivery along with its diagnostic application.

A Similar study on the usage of Nano magnetic iron oxides in drug delivery for cancer therapies has been performed by Darroudi, M. et al. [13]. The bibliographic analysis has been performed on research trends in the field. The analysis results of the study depict the progressive usage of Nano magnetic iron oxides in drug delivery. The analysis has reported the potential applications and challenges involved in the usage of these carriers along with future research directions in general multidisciplinary domains and more specifically in the treatment of colorectal cancer.

Raban et al. [14] have performed another bibliometric analysis to study the evolutionary trends of two revolutionary data analytics fields, big data, and data sciences. The study reports a significant number of publications in big data along with a continuously increasing growth in data science research publications. In addition, the results have proven a recent emergence of publications with combined applications of both big data and data science. The study evaluates the bibliometric indicators such as fields of study, journal indicators, country of origin and funding, citation indices, and authorship.

Borges do Nascimento IJ et al. [15] have performed a study for the assessment of significance for the application of big data analytics in healthcare. The study is focused on core health indicators and primacies according to the World Health Organization's General program and European program of Work. The study attempted to identify the challenges and potential opportunities of analytical tools in relation to public health. To search the data for systematic review, six databases such as Web of Science, MEDLINE, SCOPUS Embase, Cochrane Database of Systematic Reviews, and Epistemonikos have been included in the study. The core objective of the systematic reviews of the study is to assess the impact of big data analytics on people's health indicators. The study has reported that big data analytics have provided diagnosis or prediction of diseases like diabetes and mental disorders along with diagnosis or prediction of some chronic diseases with accuracy ranging from moderate to very high.

**3. Methodology.** Bibliometric methods including science mapping and performance analysis provide valuable insights about the evolution of a specific Field. [16]. Systematic Literature Review (SLR) has been used in this study to retrieve literature for the bibliometric analysis. SLR is carried out using the set of procedures for searching numerous databases using a predefined search strategy, which improves the transparency, scientific, and comprehensiveness of the study [18]. A thorough systematic literature evaluation not only improves the dependability of the research, but it also eliminates the possibility of including studies that are irrelevant [19]. The focus of this study is to explore the research output published by research studies done under the theme of 'big data and healthcare'. The study has used Scopus and Web of Science (WOS) databases to search the existing relevant literature on the topics related to the big data in healthcare. There are various reasons behind choosing the Scopus and WOS databases for retrieving the literature during this study.

1. Scopus and WOS are the most comprehensive databases available [20] [21].
2. Both databases are updated on a regular basis [22].
3. Both databases are adaptive in terms of debugging and data processing [23].

The SLR has been conducted using a precise but rigorous methodology. The search parameters and search

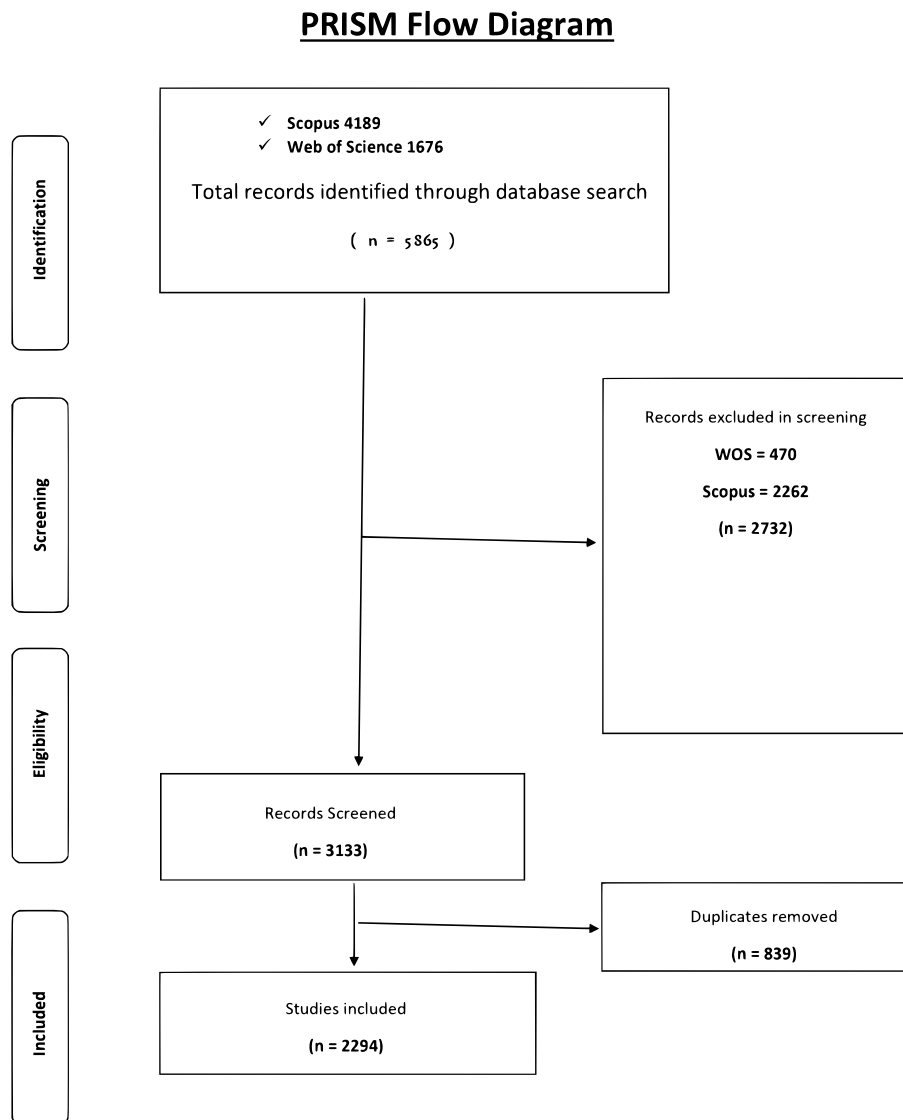


Fig. 3.1: Prism Flow Diagram of the study

refinement plays a key role in bibliometric analysis, thus needing a significant focus. Keeping this in mind, the process has been broken into three primary stages, commencing with database search in the chosen databases. In the second stage, search criteria were applied to determine which studies should be included for further analysis. Finally, bibliometric analysis has been performed to examine the scientific production of articles, and the results have been presented for performance analysis and science mapping analysis. figure 3.1 presents the prism flow diagram of the approach adopted in the study.

**3.1. Literature search.** In the literature search phase the same filtration criteria have been used for both the databases. The primary keywords such as, “big data” and “healthcare” have been used to improve the search criteria linked with big data and healthcare. While setting a search criteria the conditional params like AND, OR are used when two or more keywords are used in the query. In this study AND helper has been used in the middle of two keywords. As a result, the search terms became “big data” AND “healthcare”, implying that both terms shall be regarded equally while searching. Thus the search criteria resulted in a list

Table 3.1: Data Description

Description	Results
<b>MAIN INFORMATION ABOUT DATA</b>	
Timespan	2045:42:00
Sources (Journals, Books, etc)	1018
Documents	2294
Average years from publication	3.3
Average citations per documents	15.93
Average citations per year per doc	3.387
References	10600
<b>DOCUMENT CONTENTS</b>	
Keywords Plus (ID)	0
Author's Keywords (DE)	5465
<b>AUTHORS</b>	
Authors	7782
Author Appearances	10344
Authors of single-authored documents	196
Authors of multi-authored documents	7586
<b>AUTHORS COLLABORATION</b>	
Single-authored documents	209
Documents per Author	0.295
Authors per Document	3.39
Co-Authors per Documents	4.51
Collaboration Index	3.64

that included articles and information related to big data and healthcare. The search criteria had been further narrowed down with numerous parameters such as document type, publishing stage, and language to receive more polished and relevant data.

**3.2. Study selection.** Since the literature search in this study has been restricted to only publications present in Scopus and Web of Science databases. To ensure that the data is enough and relevant, three filtration parameters such as Document type as 'articles', stage as 'published' and Language as 'English' have been used. The main purpose of applying the language filter is to keep the focus only on a single language. The literature search resulted in 4189 articles from Scopus and 1676 articles from Web of Science during the initial keyword search. In the next step 470 articles were eliminated from Web of Science and 2262 articles from Scopus by applying the filter's parameters, leaving 2732 articles for further analysis. The information has been exported in BibTeX format from Scopus, while the data has been downloaded in plain text format from Web of Science. To ensure that the dataset records are unique, the duplicate articles have been searched, which resulted in a total of 839 duplicate articles. These duplicate documents have been dropped from the final dataset; the remaining 2294 articles have been used in further analysis of the study. The detailed description about the data used in the study has been shown in Table 3.1. Finally, the Biblioshiny a web based tool for Bibliometrix, an R based bibliometric analysis platform has been used to analyze the most significant bibliometric parameters. These parameters include the author contributions, journals, countries and institutions, exploring research hot spots, research trends and forecasting the in the selected research domain.

#### 4. Performance Analysis.

**4.1. Annual Scientific Production.** Despite the fact that big data has been there since the 1990s, it became popular when data began to rise at a rapid rate. Big data in healthcare has attracted a large number of researchers from all over the world, with an annual growth rate of 18.59 percent. The path of publication on big data in and healthcare began in 2012, as seen in Figure 4.1. Table 4.1 shows that in 2012, six research articles were published, with an average total citation of 8.33 for each publication. In 2013, the number of

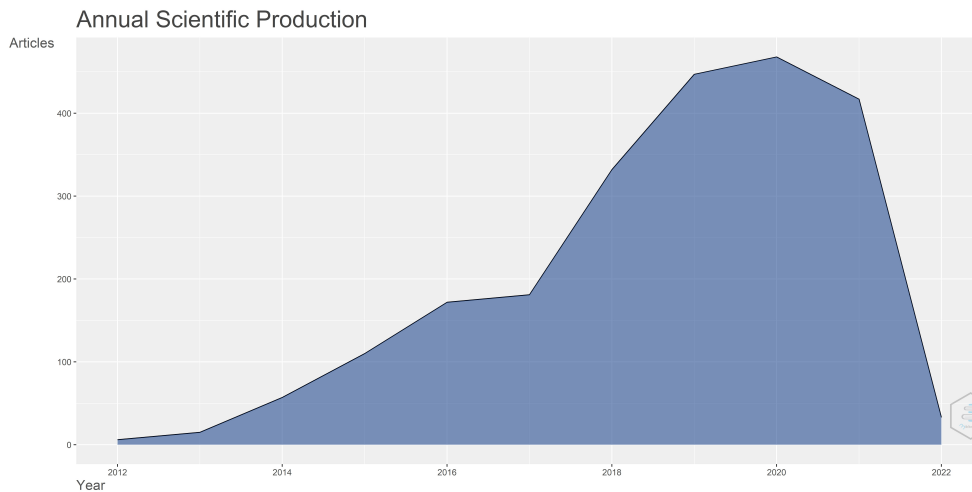


Fig. 4.1: Annual Scientific Productions

Table 4.1: Annual total citations per year

S.no.	Year	Publications	MTCperArt*	MTCperYear*	Citable Years
1	2012	6	8.333333	0.833333	10
2	2013	15	83.46667	9.274074	9
3	2014	57	26.68421	3.335526	8
4	2015	110	31.30909	4.472727	7
5	2016	172	26.47674	4.412791	6
6	2017	181	32.18232	6.436464	5
7	2018	332	24.12651	6.031627	4
8	2019	447	13.12081	4.373602	3
9	2020	468	10.50641	5.253205	2
10	2021	417	2.455635	2.455635	1
11	2022	33	0.333333		0

\*MTCperYear = Mean total citations per year MTCperArt = Mean total citations per article

studies increased to 15, and the total number of citations per article increased dramatically to 83.46, with a mean citation of 9.27. Since then, we’ve seen a significant increase with each passing year. For example, in 2014, there were 57 publications, which was three times the prior year, and we saw the same pattern in the following years, with 110 publications in 2015 and 172 publications in 2016. We did see, however, that the average total citation per article fluctuated a little and that the average total citation per year fell. In 2017, there was relatively little growth compared to 2016, with 181 publications and an average total citation per piece of 32.18. For the fourth year in a row, researchers have focused their emphasis on big in healthcare, resulting in 332, 447, 468, and 417 papers. It’s worth noting that 33 studies were published in January, indicating that 2022 will provide many publications.

**4.2. Most Globally cited documents.** According to bibliometric analysis of data from Scopus and Web Of Science (WOS), Rajkomar’s work titled ”Scalable and accurate deep learning using electronic health records” is the most cited, with 614 total citations and 122.8 total citations per year. Predictive modeling of electronic health record (EHR) data is expected to drive Customized therapy and enhance healthcare quality, according to this article published in 2018. With 433 total citations and 72.167 total citations per year, LEE JG’s study on Deep Learning in Medical Imaging remains a highly referenced publication. In 2017, LEE JG presented a study that emphasized the relevance of artificial neural networks (ANN) and the use of big data in ANN.

Table 4.2 demonstrates that articles published in 2016, 2017, and 2018 are highly referenced, indicating



Table 4.2: Authors production

S. no	Author	Paper	Total Citations	TC per Year
1	RAJKOMAR A, 2018, NPJ DIGIT MED	Scalable and accurate deep learning with electronic health records	614	122.8
2	LEE JG, 2017,	Deep Learning in Medical Imaging: General Overview	433	72.167
3	FARAHANI B, 2018,	Towards fog-driven IoT eHealth: Promises and challenges of IoT in medicine and healthcare	404	80.8
4	WANG YC, 2018, TECHNOL FORECAST SOC	Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations	376	75.2
5	DIMITROV DV, 2016,	Medical Internet of Things and Big Data in Healthcare	370	52.857
6	MOHR DC, 2013,	Behavioral Intervention Technologies: Evidence review and recommendations for future research in mental health	360	36
7	VAISHYA R, 2020,	Artificial Intelligence (AI) applications for COVID-19 pandemic	347	115.667
8	TU YF, 2020,	A Review of SARS-CoV-2 and the Ongoing Clinical Trials	345	115
9	YIN S, 2015,	Big Data for Modern Industry: Challenges and Trends	326	40.75
10	ZHANG Y, 2017, IEEE SYST J	Health-CPS: Healthcare Cyber-Physical System Assisted by Cloud and Big Data	303	50.5
11	FARAHANI B, 2018, FUTURE GENER COMP SY	Towards fog driven IoT eHealth: Promises and challenges of IoT in medicine and healthcare	302	60.4
12	CHEN M, 2017, IEEE ACCESS	Disease Prediction by Machine Learning Over Big Data from Healthcare Communities	290	48.333
13	BAKER SB, 2017, IEEE ACCESS	Internet of Things for Smart Healthcare: Technologies, Challenges, and Opportunities	255	42.5
14	HILBERT M, 2016,	Big Data for Development: A Review of Promises and Challenges	253	36.143
15	BELLE A, 2015,	Big Data Analytics in Healthcare	216	27
16	ZHONG RY, 2016, COMPUT IND ENG	Big Data for supply chain management in the service and manufacturing sectors: Challenges, opportunities, and future perspectives	215	30.714
17	FLORES M, 2013, PERS MED	P4 medicine: how systems medicine will transform the healthcare sector and society	215	21.5
18	LEE CH, 2017,	Medical big data: promise and challenges	200	33.333
19	DWIVEDI AD, 2019, SENSORS-BASEL	A Decentralized Privacy-Preserving Healthcare Block chain for IoT	199	49.75
20	MASSIE AB, 2014,	Big Data in Organ Transplantation: Registries and Administrative Claims	184	20.444

that the study is of high quality. Farahani B et al. wrote a study in 2018 called 'Towards fog-driven IoT eHealth: Promises and Challenges of IoT in Medicine and Healthcare,' which underlines the relevance of the Internet of Things (IoT) and how it may help us move from clinic-centric to patient-centric healthcare. Big Data, according to Farahani B et al, is the key to revolutionizing the healthcare environment. WANG YC's work 'Big data analytics: Understanding its capabilities and potential advantages for healthcare organizations from 2018 is one of the most referenced, with 376 total citations, followed by Dimitrov's paper from 2016 with 370 total citations. Papers by Mohr DC, Vaishya R, Tu Yf, Yin s, Yin S, Zhnag Y, and Farahani B also received 300 or more citations, with Vaishya R's work receiving the most total citations per year with 115.667. Other works have received between 180 and 300 total citations.

**4.3. Author's Production over the Time.** According to a bibliometric examination of data collected from 2012 to 2022, Zhang Y is the most productive author throughout that time span. He has 614 total citations and averages 120.8 total citations per year. Zhang Yi has over 80 articles and an h-index of 114. With around 433 total citations and 72 total citations per year, Lee H is the second most productive author of big data in healthcare, according to Fig 4.2 On the list of most productive authors, Wang Yichuan is ranked third. One of the greatest of his many writings is 'Understanding its capabilities and possible advantages for healthcare organizations.' Wang Yichuan discussed the potential benefits of big data in healthcare and how big data analytics may be used as a major pillar of the healthcare ecosystem in this study. Wang Yichuan has

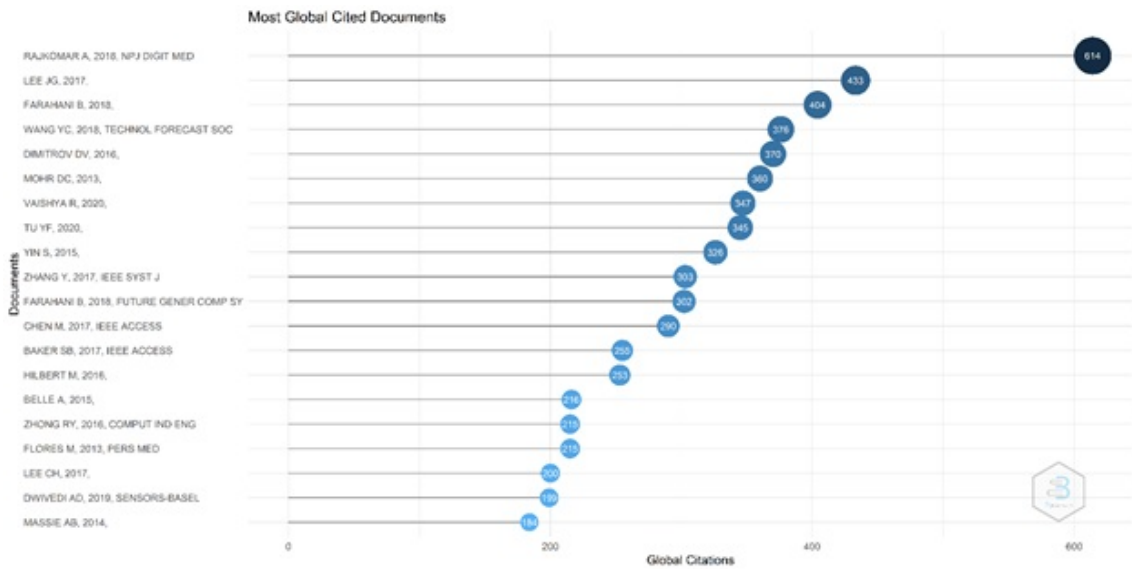


Fig. 4.2: Most globally cited documents

### Top-Authors' Production over the Time

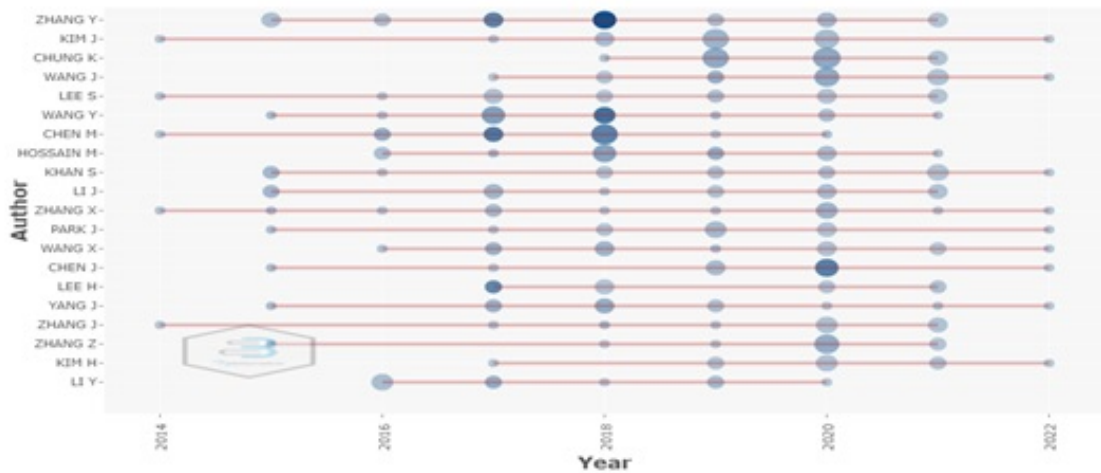


Fig. 4.3: Top-Author's Production over time

around 376 total citations, with an average of more than 75 each year. Chen M, Chen J, LI Y, Hossain M are few top authors who have contributed to the big healthcare theme.

We can refer to Fig 4.3 which highlights authors' production over the time. In this figure each article is denoted by a circular node. The size of the circular node denotes the number of the articles; the total citation is denoted by color. Darker the color means more citations. As per the figure Kim J, Lee S, Chen M and Zhang J are among few researchers who had contributed to this theme from the early period of 2014.

**4.4. The Most Productive Countries.** The distribution of publications and citations among the nations that contributed the most to the field from 2012 to 2012 will now be explored. According to the bibliometric study, the United States has contributed the most articles to the big data in healthcare issue, with 1223, which

Table 4.3: Country-wise publications and citation rank

Rank	Country	Publications	Total Citations	Average Article Citation
1	USA	1223	8987	21.552
2	CHINA	649	4791	18.788
3	INDIA	509	2804	8.987
4	UK	377	1881	15.048
5	SOUTH KOREA	244	2117	16.802
6	SPAIN	186	444	8.377
7	ITALY	165	1265	17.095
8	FRANCE	151	730	16.977
9	AUSTRALIA	141	1511	18.887
10	CANADA	141	1034	19.148
11	SAUDI ARABIA	138	864	21.073
12	GERMANY	135	566	13.163
13	NETHERLANDS	127	419	12.697
14	JAPAN	112	167	6.185
15	PAKISTAN	99	262	13.1
16	BRAZIL	52	333	12.808
17	EGYPT	52	407	16.28
18	MALAYSIA	50	691	19.194
19	PORTUGAL	47	147	8.647
20	SINGAPORE	43	311	12.958

is 100 times or more than the other nations on the list. Table 4.3 reveals that the United States has 8987 total citations and 21.552 average article citations, placing it first in both total and average article citations. China is the second country to contribute to this subject, with 649 articles with 4791 citations and an average of 18.788 citations per article. With 509 publications and 2804 total citations, India is the third most productive country in the big data in healthcare field, raising issues about the quality of the research done by Indian experts. The United Kingdom, North Korea, Spain, Italy, France, Australia, and Canada are just a handful of the many countries that have contributed to the hot issue of big data in healthcare. Fig 4.4 shows that large healthcare is something that has prompted practically every government to investigate and do study. Every country, regardless of its economy or way of living, is interested in using big data into healthcare.

The color shade in Fig 4.4 is used to help us understand each nation's contribution to this revolutionary topic. For displaying the number of publications published by a country, light blue is used as the starting shade and dark blue is used as the higher end of the scale. We can clearly distinguish the United States of America, China, and India among the darker blue colored countries, indicating that they have donated the most, while Costa Rica, Iceland, Kazakhstan, and a few other countries have contributed relatively less.

**4.5. Most Relevant Affiliations.** One of the most important components of bibliometric analysis is determining which institutions are the most prolific. Fig 4.5 shows that King Saud University is the most productive university in terms of contributions to big data in healthcare issues, with 55 papers. With 40 publications, Sanford University came in second. With 30 publications published, Harvard Medical School and Huazhong University of Science and Technology are ranked third on the list. With 28 publications published, Kyonggi University rounds out the top five. University of Florida, Oxford University, University of Minnesota, National University of Singapore, and Sejong University are among the top 10 universities that have studied big data in the healthcare field.

## 5. Science mapping analysis.

**5.1. Keyword Co-occurrence.** Key word co-occurrence analysis is widely used in knowledge mapping. Systematic literature review for knowledge mapping of current state and future growth of scientific research of any study area is very crucial, but due to its manual nature it becomes very challenging and time consuming [24]. Fig 4.3 shows the number of occurrences of the keywords in the publications. There are 793 occurrences of the word 'big data', followed by 199 occurrences of the word 'healthcare'. The occurrences of other related words

### Country Scientific Production

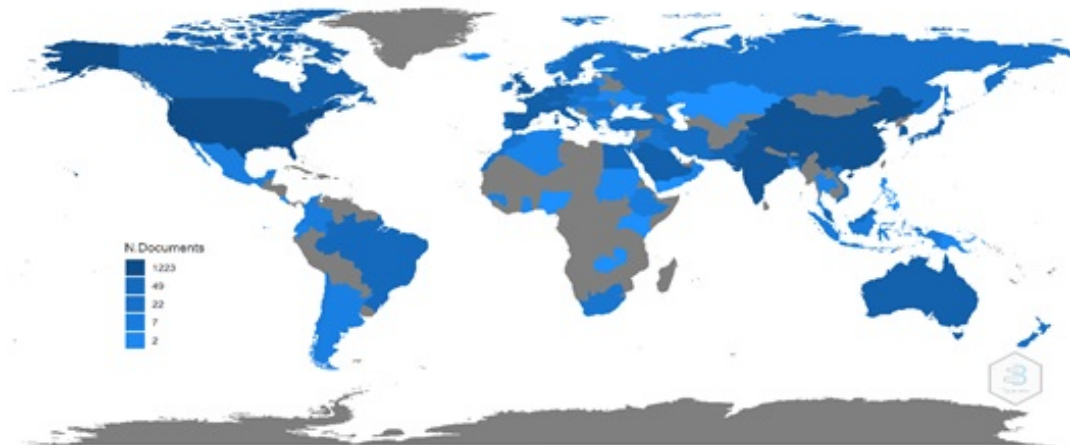


Fig. 4.4: Country-wise Scientific Production

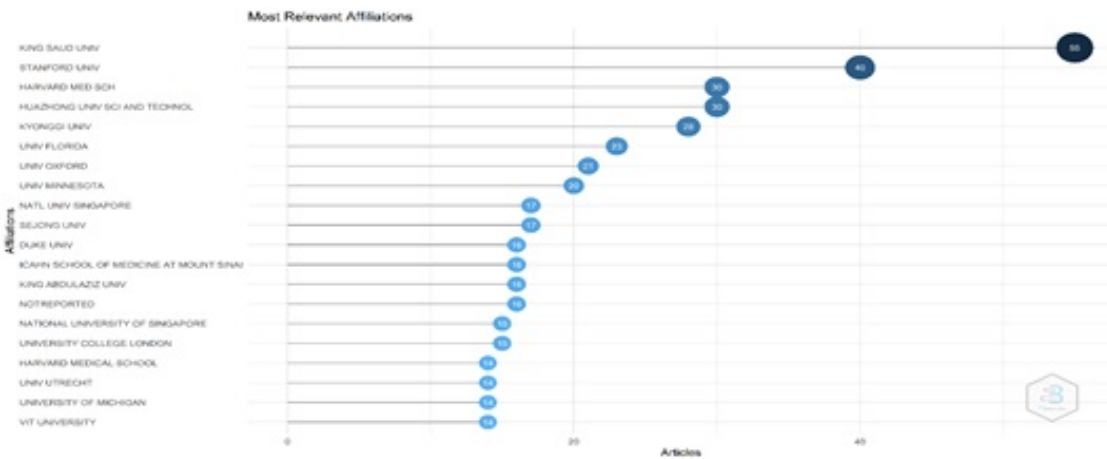


Fig. 4.5: Most relevant affiliations

like ‘machine learning’, ‘artificial intelligence’, ‘cloud computing’, ‘internet of things’ and so on are also given in the graph in decreasing order of the occurrences. These results reveal that there is a significant rise in usage and togetherness between the two words, “big data” and “healthcare” followed by the words representing the technologies very close to big data and healthcare. The word cloud of the primary and other related keywords is given in Fig 5.1 and the number of most relevant word occurrences has been shown in Fig 5.2.

. To further perform the keyword co-occurrence analysis of the key words such as ‘big data’, healthcare”, and other related keywords, a keyword co-occurrence network (KCN) has been created. KCN is represented as a graph with a set of nodes and edges. The keywords are represented as nodes and each co-occurrence is depicted by edges between the nodes. The weight of each edge connecting a pair of nodes is represented as the number of co-occurrences between the words representing the node pair [25]. The weights of the links in a KCN are visually shown as the thickness of each edge. The thickness of each edge thus is proportional to the corresponding weight of the edge. Fig 5.3 shows the KCN with three clusters represented by different colors as Red (Cluster1), Green (Cluster2), and Blue (Cluster3). The node size in the KCN represents the proportionality



Fig. 5.1: Word Cloud

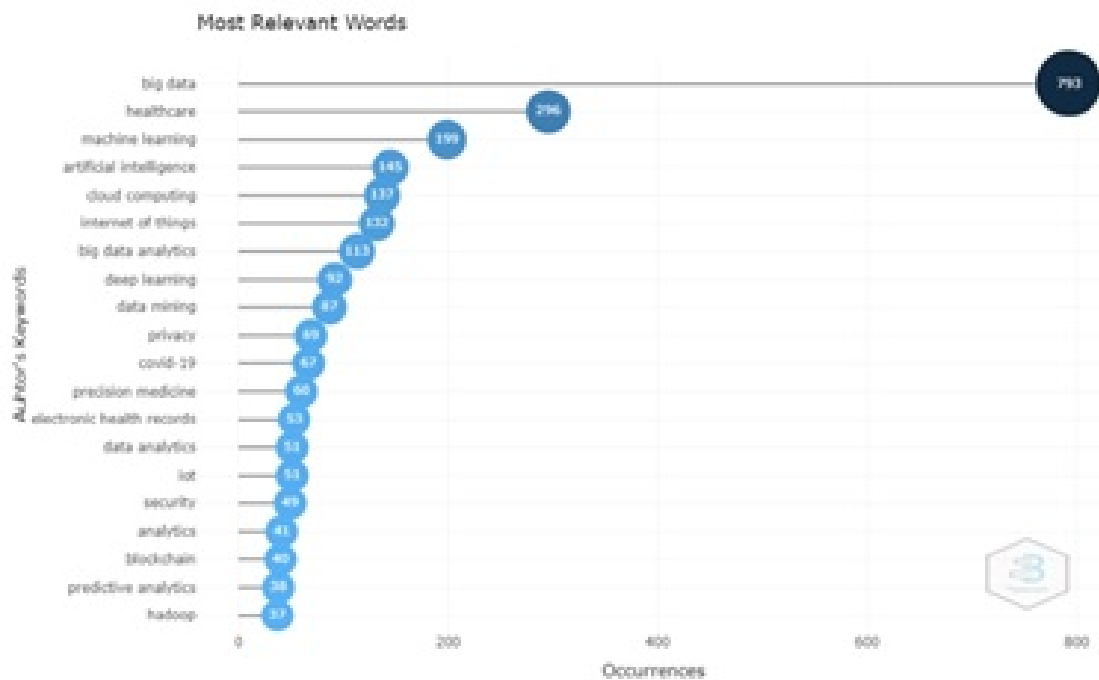


Fig. 5.2: Most Relevant Word occurrences

of frequency of words, bigger the size of node more is the occurrence of the word. As can be seen in the KCN two words in Cluster one such as, ‘big data’ followed by ‘healthcare’ are more frequent. Further the edge weight between these two words is thicker than all the other edges meaning more co-occurrences between these two words in comparison to all other word co-occurrences.

**5.2. Thematic mapping.** Thematic mapping is a conceptual structure in bibliometric. The thematic mapping outlines the conceptual structure of the keywords under consideration [26]. Thematic mapping visualizes the theme structure in the form of four quadrants of a thematic map, each quadrant representing a theme as shown in Fig 5.4. The themes are categorized in two properties such as density and centrality. Density represented by vertical axis is the degree of correlation of keywords while centrality represented by horizontal



Fig. 5.3: Keyword Co-occurrence network

axis measures the cohesiveness among the keywords. The thematic map in the figure provides the analysis of big data and healthcare. The map is illustrating the themes into four quadrants, the upper right quadrant represents motor theme, upper left quadrant represents niche theme, lower right corner represents the basic theme and lower left represents emerging or declining theme. Themes like big data, healthcare, cloud computing, AI, Machine Learning Deep learning are lying in the basic theme quadrant which are very vital in the development of the field of study. Since the thematic mapping has been done as Co-word analysis which identifies keyword clusters. They are regarded as themes, and their density and centrality being utilized to group themes and map them on a two-dimensional graph.

Themes like healthcare records, public health, data analytics and predictive analytics seen in niche quadrants depicted in the thematic map have established internal bonds. The thematic analysis thus reveals that the themes in niche quadrant such as data analytics and predictive analysis are imminent areas to be associated with big data and healthcare. Researchers in this field of study need to apply these methods to progress further in this field of study. The themes such as personal medicine, precision medicine and data science are sandwiched in all the four quadrants of the thematic map.

**5.3. Most Relevant Authors.** The results of the most relevant authors show the number of papers published by each author in the current research domain. The top 20 authors who published extensively have been shown in the results as a plot given in Fig 5.5. The numbers in the circles of the plot shows the number of publications published by each author. The results reveal that the highest number of publications, according to the Web of Science and SCOPUS database analysis; related to big data and healthcare have been contributed by ZHANG Y. The author is thus the most relevant author with 22 publications followed by KIM J with 20, which is also very close to the top author in terms of relevance of authors. CHUNG K with 19 publications in the domain is also very close. As can be seen in the plot, the least number of publications published in the studied research domain are 10 in number, which is also a significant number in terms of relevance. The results thus reveal that a good number of researchers are currently working in the research area considered in this study. Further the results of Lotka's law analysis, given in Fig 5.5, which gives the frequency of publication by the relevant authors in big data and healthcare.

## 6. Annual Source Growth.

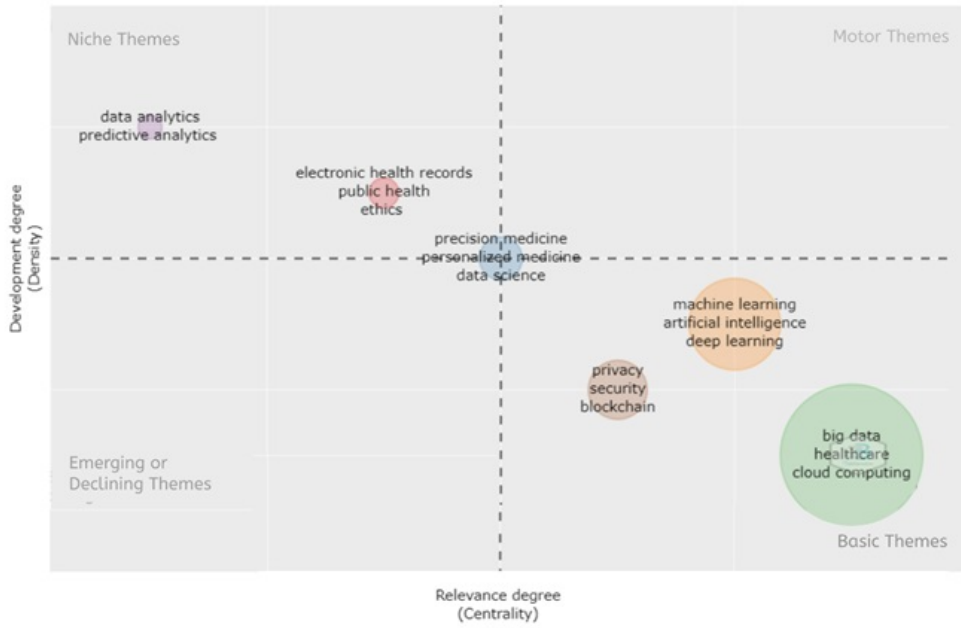


Fig. 5.4: Thematic Mapping

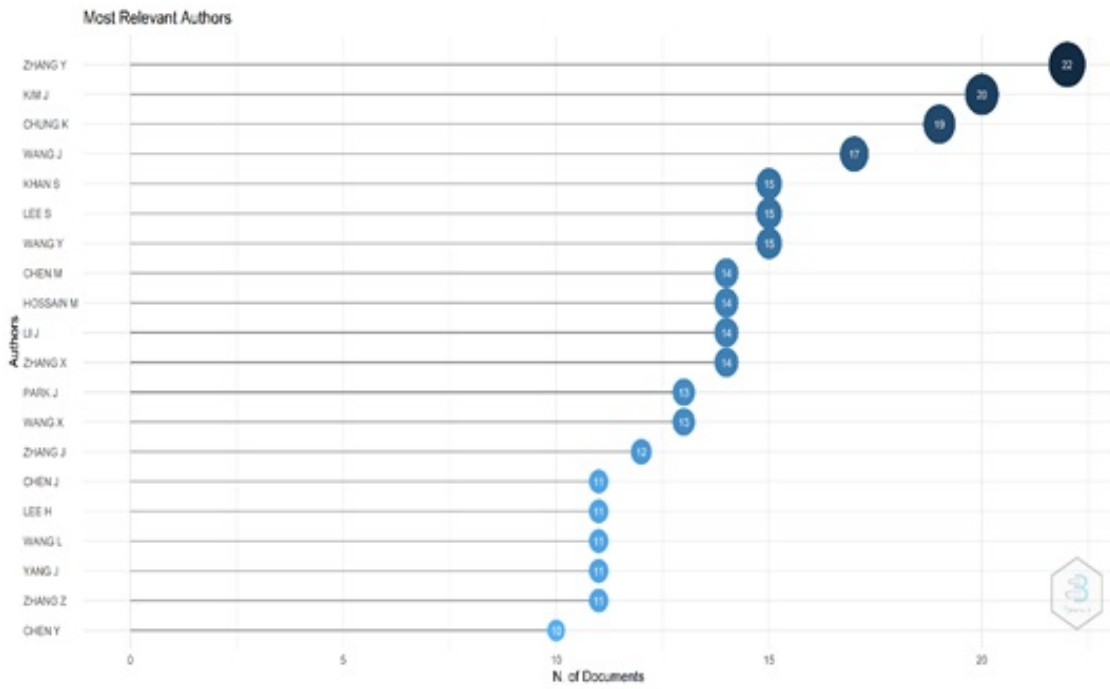


Fig. 5.5: Most Relevant Authors

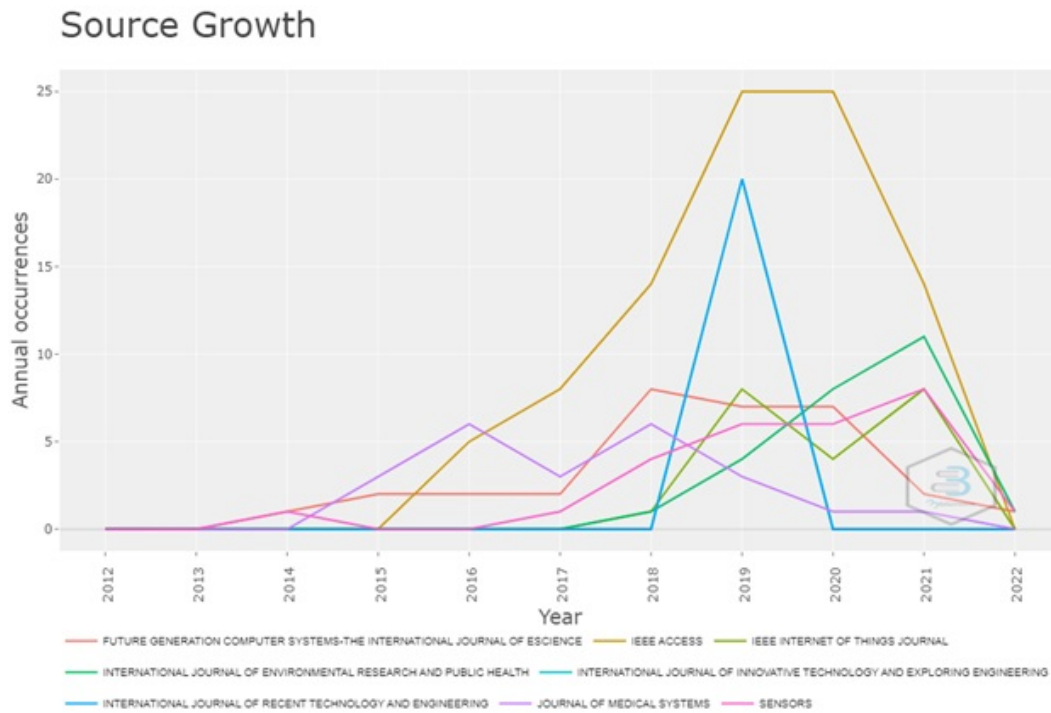


Fig. 6.1: Annual Source growth

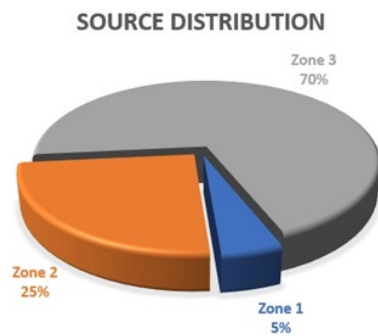


Fig. 6.2: Source distribution percentage

**6.1. Annual Source Distribution.** The annual source growth is another important bibliometric indicator analyzed in the study. This indicator depicts the growth distribution of publications related to the given domain by source. The results shown in Fig 6.1 depict the number of publications available in top 8 journals. The overall results reveal 2294 articles published by 1019 journal sources. According to the Bradford’s Law analysis there are three zones: Zone1 includes 49 journals with 758 published articles (33.04%), Zone2 consisted of 252 journals, with 779 published articles (33.95%), and Zone3 consists 718 journals with 757 articles (32.99%). The Bradford analysis graph for top 50 sources is shown in Fig 6.3 the analysis data for the graph has been depicted in Table 6.1. The source distribution percentage has been illustrated in Fig 6.2.

**6.2. Collaboration Network Analysis.** Collaboration Networks are used to visualize the research collaborations between authors, research groups, institutions, and countries [27] [28]. Collaboration networks can



Table 6.1: Source clustering through Bradford's Law

Rank	SO	Freq	cumFreq
1	IEEE ACCESS	91	91
2	FUTURE GENERATION COMPUTER SYSTEMS-THE INTERNATIONAL JOURNAL OF ESCIENCE	32	123
3	SENSORS	27	150
4	INTERNATIONAL JOURNAL OF ENVIRONMENTAL RESEARCH AND PUBLIC HEALTH	25	175
5	JOURNAL OF MEDICAL SYSTEMS	23	198
6	IEEE INTERNET OF THINGS JOURNAL	21	219
7	INTERNATIONAL JOURNAL OF INNOVATIVE TECHNOLOGY AND EXPLORING ENGINEERING	20	239
8	INTERNATIONAL JOURNAL OF RECENT TECHNOLOGY AND ENGINEERING	20	259
9	JOURNAL OF ADVANCED RESEARCH IN DYNAMICAL AND CONTROL SYSTEMS	19	278
10	YEARBOOK OF MEDICAL INFORMATICS	19	297
11	BMC MEDICAL INFORMATICS AND DECISION MAKING	18	315
12	JOURNAL OF BIG DATA	18	333
13	INTERNATIONAL JOURNAL OF MEDICAL INFORMATICS	17	350
14	TECHNOLOGICAL FORECASTING AND SOCIAL CHANGE	17	367
15	JOURNAL OF AMBIENT INTELLIGENCE AND HUMANIZED COMPUTING	15	382
16	PLOS ONE	15	397
17	SUSTAINABILITY	15	412
18	BIG DATA	14	426
19	INTERNATIONAL JOURNAL OF ENGINEERING AND TECHNOLOGY(UAE)	14	440
20	JOURNAL OF COMPUTATIONAL AND THEORETICAL NANOSCIENCE	14	454
21	JOURNAL OF HEALTHCARE ENGINEERING	14	468
22	BMJ OPEN	13	481
23	JOURNAL OF THE AMERICAN MEDICAL INFORMATICS ASSOCIATION	13	494
24	APPLIED SCIENCES-BASEL	12	506
25	HEALTH AND TECHNOLOGY	12	518
26	HEALTHCARE INFORMATICS RESEARCH	12	530
27	IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS	12	542
28	INTERNATIONAL JOURNAL OF ENGINEERING AND ADVANCED TECHNOLOGY	12	554
29	INTERNATIONAL JOURNAL OF INFORMATION MANAGEMENT	12	566
30	JOURNAL OF SUPERCOMPUTING	12	578
31	BIG DATA RESEARCH	11	589
32	CLUSTER COMPUTING-THE JOURNAL OF NETWORKS SOFTWARE TOOLS AND APPLICATIONS	11	600
33	JOURNAL OF BIOMEDICAL INFORMATICS	11	611
34	MULTIMEDIA TOOLS AND APPLICATIONS	11	622
35	NPJ DIGITAL MEDICINE	11	633
36	WIRELESS PERSONAL COMMUNICATIONS	11	644
37	COMPUTER METHODS AND PROGRAMS IN BIOMEDICINE	10	654
38	INTERNATIONAL JOURNAL OF SCIENTIFIC AND TECHNOLOGY RESEARCH	10	664
39	CIN-COMPUTERS INFORMATICS NURSING	9	673
40	FUTURE GENERATION COMPUTER SYSTEMS	9	682
41	IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS	9	691
42	INFORMATION SYSTEMS FRONTIERS	9	700
43	INTERNATIONAL JOURNAL OF ADVANCED COMPUTER SCIENCE AND APPLICATIONS	9	709
44	INTERNATIONAL JOURNAL OF APPLIED ENGINEERING RESEARCH	9	718
45	APPLIED SCIENCES (SWITZERLAND)	8	726
46	HEALTH INFORMATICS JOURNAL	8	734
47	INTERNATIONAL JOURNAL OF ADVANCED SCIENCE AND TECHNOLOGY	8	742
48	INTERNATIONAL JOURNAL OF HEALTHCARE MANAGEMENT	8	750
49	JOURNAL OF MEDICAL IMAGING AND HEALTH INFORMATICS	8	758
50	PERSONALIZED MEDICINE	8	766



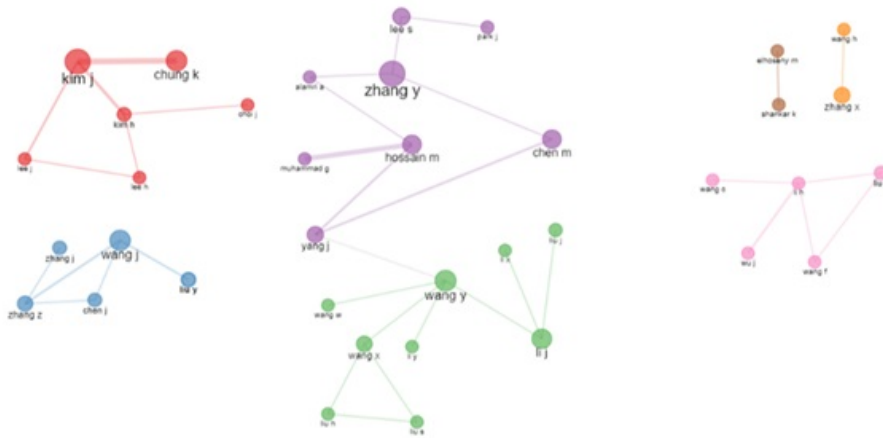


Fig. 6.4: Author Collaborations Network

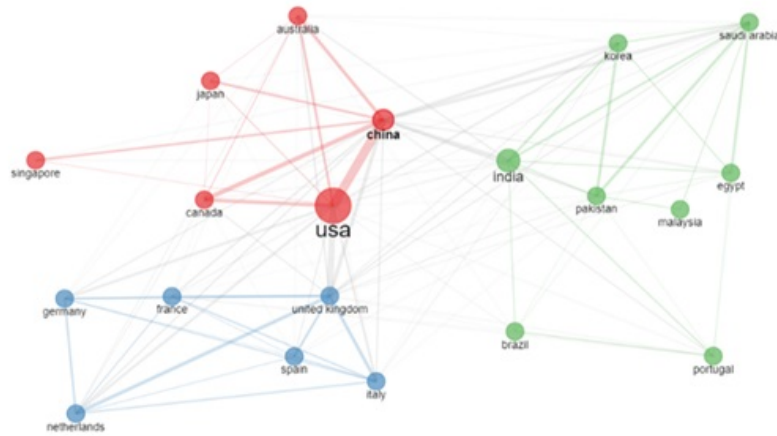


Fig. 6.5: Country Collaborations Network

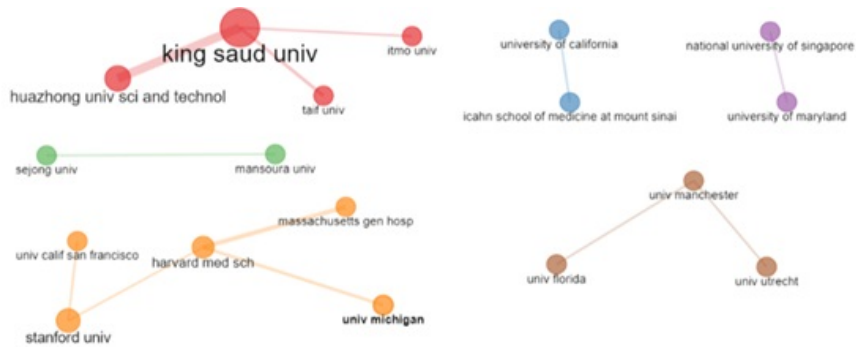


Fig. 6.6: Institutional Collaborations Network

during the past decade which led to the exponential growth of the number of research studies in the field [29][30]. This study is focused on analysis of many crucial bibliometric indicators in an interdisciplinary research domain such as big data and healthcare. One of the contributions of this study is to identify the research studies and their reported outputs in the field of big data in healthcare. The study will be used as a reference by the research community, editors of research journals, industry professionals and academicians to recognize the recent state of scientific research, most prominent research articles, most prolific researchers, most prominent sources and potential collaborations between authors, institutions and countries in the field. The study reviews and summarizes the scientific literature, identifies the future research trends and directions.

A total of 5685 articles were extracted from the Scopus and web of science databases out of which 2294 were selected for further analysis in this study. The study contributes prominently to the body of research. The results of the study reveal that the most relevant publishing source was IEEE ACCESS. Further, the analysis results reveal that ZHANG Y, KIM J. are the most relevant authors with the highest number of publications in the study domain. Similarly, the most relevant country which contributed to the research field over the last decade in terms of number of publications is the USA. In the institutional contribution analysis, King Saud University is the most productive university in terms of contributions to big data in healthcare followed by Sanford University. In the thematic mapping, the results reveal that the big data in healthcare is significantly emerging along with the themes like, data analytics, predictive analytics, electronic health records, public health and are closely connected to the field of big data in healthcare.

**Acknowledgments.** Authors would like to thank their friends, colleagues and their respective institutions for support in conducting in this study.

#### REFERENCES

- [1] Dash, S., Shakyawar, S.K., Sharma, M. et al. Big data in healthcare: management, analysis and future prospects. *J Big Data* 6, 54 (2019). <https://doi.org/10.1186/s40537-019-0217-0>
- [2] IDC , The Digitization of the World From Edge to Core <https://www.idc.com/getdoc.jsp?containerId=prUS47560321>, accessed on : 09 Feb 2022
- [3] Aijaz Ahmad Reshi, Furqan Rustam, Arif Mehmood, Abdulaziz Alhossan, Ziyad Alrabiah, Ajaz Ahmad, Hessa Alsuwailam, Gyu Sang Choi, An Efficient CNN Model for COVID-19 Disease Detection Based on X-Ray Image Classification, *Complexity*, vol. 2021, Article ID 6621607, 12 pages, 2021. <https://doi.org/10.1155/2021/6621607>
- [4] Reshi AA, Ashraf I, Rustam F, Shahzad HF, Mehmood A, Choi GS. Diagnosis of vertebral column pathologies using concatenated resampling with machine learning algorithms. *PeerJ Computer Science* 7:e547 <https://doi.org/10.7717/peerj-cs.547>
- [5] Aria and Cuccurullo, (2017), A bibliometric analysis of pandemic and epidemic studies in economics: future agenda for COVID-19 research, <https://www.sciencedirect.com/science/article/pii/S2590291121000619#bib2>
- [6] Guandong Song, Jiyang Wu, Sihui Wang, Text Mining in Management Research: A Bibliometric Analysis”, *Security and Communication Networks*, vol. 2021, Article ID 2270276, 15 pages, 2021. <https://doi.org/10.1155/2021/2270276>
- [7] Guo, Yuqi, et al. Artificial Intelligence in Health Care: Bibliometric Analysis.” *Journal of medical Internet research* vol. 22,7 e18228. 29 Jul. 2020, doi:10.2196/18228
- [8] Dash, S., Shakyawar, S.K., Sharma, M. et al. Big data in healthcare: management, analysis, and prospects. *J Big Data* 6, 54 (2019). <https://doi.org/10.1186/s40537-019-0217-0>
- [9] Galetsi, P. and Katsaliaki, K. (2020), Big data analytics in health: an overview and bibliometric study of research activity. *Health Info Libr J*, 37: 5-25. <https://doi.org/10.1111/hir.12286>
- [10] Hu, Yuanzhang MDa; Yu, Zeyun MDB; Cheng, Xiaoen MDa,\*; Luo, Yue MDa; Wen, Chuanbiao MDa,\* A bibliometric analysis and visualization of medical data mining research, *Medicine*: May 29, 2020 - Volume 99 - Issue 22 - p e20338 DOI: 10.1097//MD.00000000000020338
- [11] Ale Ebrahim S., Zamani Pedram M., Ale Ebrahim N. (2020) Current Status of Systemic Drug Delivery Research: A Bibliometric Study. In: Lai WF. (eds) *Systemic Delivery Technologies in Anti-Aging Medicine: Methods and Applications. Healthy Ageing and Longevity*, vol 13. Springer, Cham. <https://doi.org/10.1007/978-3-030-54490-4-2>
- [12] Wu Haiyang, Tong Linjian, Wang Yulin, Yan Hua, Sun Zhiming, Bibliometric Analysis of Global Research Trends on Ultrasound Microbubble: A Quickly Developing Field”, *Frontiers in Pharmacology*, 2021, <https://www.frontiersin.org/article/10.3389/fphar.2021.646626>
- [13] Darroudi, M., Gholami, M., Rezayi, M. et al. An overview and bibliometric analysis on the colorectal cancer therapy by magnetic functionalized nanoparticles for the responsive and targeted drug delivery. *J Nanobiotechnol* 19, 399 (2021). <https://doi.org/10.1186/s12951-021-01150-6>
- [14] Raban, D.R., Gordon, A. The evolution of data science and big data research: A bibliometric analysis. *Scientometrics* 122, 1563–1581 (2020). <https://doi.org/10.1007/s11192-020-03371-2>
- [15] Borges do Nascimento IJ, Marcolino MS, Abdulazeem HM, Weerasekara I, Azzopardi-Muscat N, Gonçalves MA, Novillo-Ortiz

- D Impact of Big Data Analytics on People's Health: Overview of Systematic Reviews and Recommendations for Future Studies *J Med Internet Res* 2021;23(4): e27275
- [16] Cobo MJ, López-Herrera AG, Herrera-Viedma E, Herrera F (2011) Science mapping software tools: review, analysis, and cooperative study among tools. *J Am Soc Inform Sci Technol* 62:1382–1402. <https://doi.org/10.1002/asi.21525>
- [17] Pittway, L. (2008) Systematic literature reviews. In Thorpe, R., Holt, R. The SAGE dictionary of qualitative management research. London: SAGE Publications.
- [18] Tang, 2019, A systematic literature review and analysis on mobile apps in m-commerce: Implications for future research, *Electronic Commerce Research and Applications* doi:10.1016/j.elerap.2019.100885
- [19] Mugomeri E, Bekele BS, Mafaesa M, et al (2017) A 30-year bibliometric analysis of research coverage on HIV and AIDS in Lesotho. *Health Res Policy Syst* 15:21. <https://health-policy-systems.biomedcentral.com/articles/10.1186/s12961-017-0183-y>
- [20] Zyoud SH, Fuchs-Hanusch D (2017) A bibliometric-based survey on AHP and TOPSIS techniques. *Expert Syst Appl* 78:158–181. <https://doi.org/10.1016/j.eswa.2017.02.016>
- [21] Borrett SR, Sheble L, Moody J, Anway EC (2018) Bibliometric review of ecological network analysis: 2010–2016. *Ecol Model* 382:63–82. <https://doi.org/10.1016/j.ecolmodel.2018.04.020>
- [22] Aznar-Sánchez JA, Velasco-Muñoz JF, Belmonte-Ureña LJ, Manzano-Agugliaro F (2019) Innovation and technology for sustainable mining activity: a worldwide research assessment. *J Clean Prod* 221:38–54. <https://doi.org/10.1016/j.jclepro.2019.02.243>
- [23] Radhakrishnan S, Erbis S, Isaacs JA, Kamarthi S (2017) Novel keyword co-occurrence network-based methods to foster systematic reviews of scientific literature. *PLoS ONE* 12(3): e0172778. <https://doi.org/10.1371/journal.pone.0172778>
- [24] Lozano, S., Calzada-Infante, L., Adenso-Díaz, B. et al. Complex network analysis of keywords co-occurrence in the recent efficiency analysis literature. *Scientometrics* 120, 609–629 (2019). <https://doi.org/10.1007/s11192-019-03132-w>
- [25] Della Corte, V.; Del Gaudio, G.; Sepe, F.; Sciarelli, F. Sustainable Tourism in the Open Innovation Realm: A Bibliometric Analysis. *Sustainability* 2019, 11, 6114. <https://doi.org/10.3390/su11216114>
- [26] Wu, W., Xie, Y., Liu, X., Gu, Y., Zhang, Y., Tu, X., Tan, X. (2019). Analysis of Scientific Collaboration Networks among Authors, Institutions, and Countries Studying Adolescent Myopia Prevention and Control: A Review Article. *Iranian journal of public health*, 48(4), 621–631
- [27] A. A. Reshi, A. Alsaeedi and S. Shafi, Development and Web Performance Evaluation of Internet of Things testbed, 2019 International Conference on Computer and Information Sciences (ICIS), Sakaka, Saudi Arabia, 2019, pp. 1-6, doi: 10.1109/ICISCI.2019.8716436.
- [28] Gubbi J, et al. Internet of Things (IoT): a vision, architectural elements, and future directions. *Future Gener Comput Syst.* 2013;29(7):1645–60, <https://doi.org/10.1016%2Fj.future.2013.01.010>
- [29] Batko, K., Ślęzak, A. The use of Big Data Analytics in healthcare. *J Big Data* 9, 3 (2022). <https://doi.org/10.1186/s40537-021-00553-4>
- [30] Aria, M.; Cuccurullo, C.; D'Aniello, L.; Misuraca, M.; Spano, M. Thematic Analysis as a New Culturomic Tool: The Social Media Coverage on COVID-19 Pandemic in Italy. *Sustainability* 2022, 14, 3643. <https://doi.org/10.3390/su14063643>
- [31] Reshi AA, Rustam F, Aljedaani W, Shafi S, Alhossan A, Alrabiah Z, Ahmad A, Alsuwailem H, Almangour TA, Alshammari MA, Lee E, Ashraf I. COVID-19 Vaccination-Related Sentiments Analysis: A Case Study Using Worldwide Twitter Dataset. *Healthcare*. 2022
- [32] Swain, B. K., Khan, M. Z., Chowdhary, C. L., Alsaeedi, A. SRC: Superior Robustness of COVID-19 Detection from Noisy Cough Data Using GFCC, *Computer Systems Science and Engineering* 2023, 46(2), 2337-2349
- [33] Javaid, A., Siddique, M.A., Reshi, A.A. et al. Coal mining accident causes classification using voting-based hybrid classifier (VHC). *J Ambient Intell Human Comput* (2022)
- [34] F. Rustam et al., COVID-19 Future Forecasting Using Supervised Machine Learning Models, in *IEEE Access*, vol. 8, pp. 101489-101499, 2020
- [35] Khan, Y. F., Kaushik, B., Chowdhary, C. L., Srivastava, G. (2022). Ensemble Model for Diagnostic Classification of Alzheimer's Disease Based on Brain Anatomical Magnetic Resonance Imaging. *Diagnostics*, 12(12), 3193
- [36] F. Rustam et al., Sensor-Based Human Activity Recognition Using Deep Stacked Multilayered Perceptron Model, in *IEEE Access*, vol. 8, pp. 218898-218910, 2020
- [37] Shabana Shafi , Aijaz Ahmad Reshi and A. Kumaravel, "Wireless Sensor Network based Early Warning and Alert System for Radioactive Radiation Leakage", *Middle-East Journal of Scientific Research* 19 (12): 1602-1608, 2014
- [38] Padinjappurathu Gopalan, S., Chowdhary, C. L., Iwendi, C., Farid, M. A., Ramasamy, L. K. (2022). An efficient and privacy-preserving scheme for disease prediction in modern healthcare systems. *Sensors*, 22(15), 5574

*Edited by:* Chiranji Lal Chowdhary

*Special Issue on:* Scalable Machine Learning for Health Care: Innovations and Applications

*Received:* Mar 19, 2023

*Accepted:* May 28, 2023





## EFFICIENT NET-BASED TRANSFER LEARNING TECHNIQUE FOR FACIAL AUTISM DETECTION

**Abstract.** Autism Spectrum Disorder is a neurological disorder in which an individual faces life-long effects in communication and interaction with others. Nowadays, the Autism Spectrum disorder ratio is increasing drastically more than ever before. Autism can be identified at all developmental levels as a "behavioural condition," and its symptoms often arise between the ages of two and four. The ASD issue starts during puberty and persists through adolescence and adulthood. Children with ASD use both nonverbal and verbal behaviour to communicate, and they struggle with joint attention and social reciprocity. Children with autism are frequently socially isolated as a result of these problems. Through very expensive and time-consuming screening exams, autism spectrum features can be identified. As one of the possible mirrors of the brain, children's faces can be utilised as a biomarker and as a quick and convenient technique for the early identification of ASD. An effective, genuine, and automatic method of face-based spectrum disorder identification is required. In this study we compare the transfer learning approach used for autism identification with the convolutional neural network (CNN)-based efficient-net strategy to identify autistic children using facial images. We used an open-source Kaggle dataset and evaluated the model performance in terms of accuracy, confusion matrix, precision, recall, and F1 measure. Efficient shows an accuracy of 97% on the benchmark dataset and beats the baseline technique of transfer learning-based approaches. This study can be used to help medical professionals validate their initial screening procedures for finding youngsters with ASD disease.

**Key words:** Autism spectrum disorder, Convolutional Neural Network (CNN), Transfer Learning, EfficientNet, Facial Images

**1. Introduction.** Autism Spectrum Disorders (ASD) are rapidly increasing in all age groups of the population today. However, it is the most alarming developmental disorder that disrupts children's social skills, communication skills, and imagination. Children with autism have problems with repetitive behaviour patterns, anxiety, self-harm, abnormal sleep patterns, behaviour-altering aggression, and attention deficits. Autism symptoms vary in children from mild to severe. An early diagnosis and detection can help treat this autistic problem. The medical team, with the help of parents and guardians, manages the ASD screening instruments, which are the assessments used to identify autism spectrum disorder. Despite deep research, the ASD neural mechanism problem is still unclear. Generally, ASD diagnosis is based on behaviour, not on the cause or mechanism [13]. In ASD diagnosis, Genetic tests are widely used to identify genetic causes, however, genetic tests provide only indication about potential risk and not diagnosis information. The researchers are using brain imaging techniques to diagnose ASD [7]. However, imaging-based techniques have some limitations in terms of datasets. Researchers are investigating machine learning methods to evaluate ASD swiftly and economically [28]. Machine learning-based solutions offer quick, accurate, and automatic procedures for ASD detection. Machine learning, a branch of artificial intelligence, has the potential to improve neurological disease identification using computer approaches [24, 14]. There are different ways to detect ASD based on different modalities of neuroimaging data like Electroencephalography (EEG), Magnetoencephalography (MEG), Electrocorticography (ECoG), Magnetic resonance imaging (MRI) and Functional near-infrared spectroscopy (fNIRS) [9].

**1.1. Motivation.** ASD (ASD) is a condition marked by social difficulties. In accordance with the social motivation theory, ASD causes a decline in social motivation because affected individuals find social stimuli less satisfying than neurotypical individuals. The social motivation hypothesis offers a developmental perspective on how social deficiencies in ASD may later manifest as faulty reward processing. According to social theory, young ASD patients pay less attention from an early age to social cues like faces and gaze direction. Due to the lack of possibilities for social learning (such as friendships, cooperative play, and joint attention), the growth of social skills is hampered as a result. The social motivation theory identifies impaired social approach and involvement as two crucial diagnostic features of ASD. However, these mentioned procedures are expensive and are out of reach treatment, especially in developing countries. Late diagnosis and identification ASD is

commonly due to:

- The identification of ASD children near the age of two is done through interactive sessions that call for clinical professionals [23].
- The lack of availability of the appropriate physicians, especially in underdeveloped countries [27].
- The people who are not aware of the ASD disorder do not take this disease as seriously as they should at an early stage.
- Additionally, due to the high expenses of the sophisticated equipment and qualified staff needed for these tests, Children from racial and ethnic minorities who get a primary screening are less likely to undergo further medical exams [30].

It is clear from the above discussion that the current developments in the field of deep learning, particularly convolutional neural networks, we can rely on automatic feature extraction. We proposed simple and lightweight models that provide better results than transfer learning-based techniques. Researchers are investigating and creating novel diagnostic methods for the early detection of ASD based on facial expression as a new area of study. The distinctive characteristics of an ASD person makes them easier to recognize from facial expressions. According to studies conducted at the University of Missouri, children who have autism tend to have particular facial characteristics, such as a big upper face and wide-set eyes. Compared to children without the condition, their faces frequently have a shorter centre section, encompassing the cheeks and nose [3]

- Our proposed technique provides better results with simple models.
- The proposed technique beat the state of art transfer learning techniques and we make comparisons to prove the effectiveness of our technique.

The remaining part of paper is organized in sections. Section 2 is about the state of art literature review, section 3 is about proposed methodology. In section 4 we discuss about performance metrics and results and the final section, section 5, is about conclusion and future work.

**2. Literature Review.** The autism condition is associated with brain development issues affecting how individuals interact and communicate. Marotta et al. [18] define autism as "Complex neurobehavioral and neurodevelopmental conditions characterized by altered sensory processing, restrained and repetitive patterns of behaviour or interests, and impaired social interaction and communication." Children affected by autism often portray limited and monotonous patterns of conduct. The disorder is not easily diagnosed because it affects people differently, making it hard to distinguish the unique features that characterise the condition. The term 'spectrum' depicts the wide range of signs and symptoms that affect patients with varying severity. Autism spectrum disorder is often detected during early childhood and impacts how people function in society. For instance, some people are easily irritated and have difficulty maintaining relationships, while others suffer from attention deficit, affecting their education adversely. It is impossible to ascertain the causes of autism disease though some cases are associated with genetic inheritance, preterm births, or head injuries at a young age. Khodatars et al. [15] argue that autism can be diagnosed using numerous diagnostic protocols, especially after children mark their third birthday. The ailment cannot be cured, but early diagnosis, treatment, and intensive management help improve the lives of the affected individuals. Hence, autism is a lifelong-manageable condition though its prevalence over the last decade raises global concern.

Over the last decade, the cases of autism spectrum disorder have increased at a concerning rate. Chiarotti et al.[8] state autism spectrum disorder (ASD) frequency has significantly increased over the past few decades, leading to allegations that autism is an "epidemic." The cases reported in different regions indicate that the incidents diagnosed are increasing. However, some may argue that the diagnosis procedures have improved over the years with proper recording, unlike in the past. For instance, countries in the global south report more cases than in the past due to enhanced capability to track and diagnose the ailment. Contemporary, approximately one in 54 children is suffering from ASD globally. The condition is more prevalent among boys than in girls at the rate of 3.68% and 1.25%, respectively. The cases might still be more than the current studies report since some parents do not like to disclose their children's struggles due to fear of stigmatisation. Botha et al.[6] states that the media and some societies frame autism negatively and stereotype autism in a way that encourages individuals suffering from the condition to conceal or camouflage as non-autistic even when it impacts their functioning and wellness. In the US, autism affects about 2% of all children, creating a heavy financial burden on taxpayers as the affected children need significant medical, educational, and social support. As a result,



the growing number of ASD cases discovered using various diagnostic techniques affects many individuals and families.

Mottron et al.[21] states that healthcare professionals rely on several diagnostic approaches to identify children with autism where behavioural observation remains critical but artificial intelligence might revolutionise ASD screening and diagnosis. According to this study, "The planning of intervention and educational services is poorly aided by a single categorical diagnosis, which encompasses such heterogeneity of developmental history, intelligence, comorbidity, and severity." Physicians do not have one verified diagnostic approach recommended for use but depend on varying ones, complicating the process. A child might be misdiagnosed as autistic due to less severe symptoms associated with autism, while others fail to be identified until later. McCarty et al.[19] defined ASD, categorised, and diagnosed per the criteria established by DSM-5, though the manual does not provide a formal test as with other disorders. Therefore, most medics depend on behavioural measurements to evaluate suspected cases.

Learning based approach have great contribution in the prediction of autism disease [5],[29]. Parisot et al.[25] built a population graph using a graph-based technique and trained a graph convolution network (GCN) to do so. On the ABIDE1 dataset, they attained a classification accuracy of 70.4%. Haque et al.[10] developed "deep convolutional neural network" deep learning methodologies, which drew inspiration from the VGGNet network family. Utilising the well-known FER2013 dataset [2], a Deep Convolutional Neural Network (DCNN) algorithm was trained. The brightness of the images in this dataset has been altered to test the model's performance in various lighting scenarios and to recognize the expressions on the faces of children with autism. Li et al. [17] used an end-to-end learning-based strategy for identifying autism spectrum disorder (ASD) using facial data such as expressions, action units, arousal, and valence. This study uses convolutional neural network representations of several facial traits that have been trained on real-world pictures. The video dataset contains 105 children (62 with autism spectrum disorders and 43 without). Ahmed et al. [3] classified ASD subjects from healthy controls using a support vector machine (SVM) and a limited Restricted Boltzmann Machine (RBM) to extract characteristics from fMRI data. The dataset is initially pre-processed, which includes slice time correction and normalization. This study employed 105 Typical control (TC) and 79 ASD patients from the renowned database ABIDE. The results demonstrate that when classifying ASD using grid-search cross-validation, the proposed framework performs very well. The outcomes also imply that merging RBM and SVM techniques may be employed as an ASD detection tool in the future. Khosla et al.[16] used transfer learning-based model of MobileNet, InceptionV3, and InceptionResNetV2 models, and reported poorer accuracy compared to other studies. They also used the MD5 hash technique to remove duplicates from the dataset. Mishra et al.[20] exploited surface morphometric properties of T1-weighted structural magnetic resonance imaging (sMRI) to develop a machine-learning approach for the detection of ASD. The proposed study integrates classification evaluation of the used machine learning models based on the surface morphometric characteristics of the left and right hemispheres of the brain. The Decision Tree (DT) and Random Forest (RF) are employed for categorisation.

Hosseini et al.[11] has good contribution in identification of autism and used CNN-based architecture MobileNet as transfer learning technique. The proposed method has a good accuracy result on images of small children. The visual features were extracted using the pre-trained deep learning models, which used three fully linked layers topped by a dense layer to forecast. Rahman et al.[22] used CNN-based architecture XceptionNet which was utilised to identify the autistic disease. On the facial image dataset for autism identification, Alsaade et al. [4] used three CNN-based architectures: Xception, VGG19, and NASNETMobile, and achieved the maximum accuracy. All the aforementioned CNN-based models, which were intensively trained on the ImageNet dataset, which contains 14 million images divided into 1000 categories, are used to extract attributes from the photographs in the Kaggle autistic image dataset (<https://www.kaggle.com/general/123978>). Using fMRI scans from the ABIDE-1 dataset, Prasad et al.[26] classified Using a multilayer perceptron (MLP) based classification model with autoencoder pretraining, ASD is distinguished from Typically Developing (TD). The suggested method identified the correlations between brain regions that contribute most to the categorization problem using the Integrated Gradients (IG) and DeepLIFT approaches. Following regions are shown to be related with this analysis: left lingual gyrus, right insula lobe, right cuneus, right middle frontal gyrus, and left superior temporal gyrus.

Table 2.1: Summary of Literature Review

Study	Year	Technique	Dataset	Performance Metrics
Parisot et al. [25]	2018	Graph convolution network (GCN)	ABIDE and ADNI	Accuracy and AUC
Haque et al. [10]	2018	Deep convolutional neural network (DCNN)	Fer2013	Accuracy Score
Li et al. [17]	2019	Convolution neural network (CNN)	Video dataset	F1 score
Ahmed et al. [3]	2020	Restricted boltzmann machines with SVM	ABIDE-I	Accuracy, ROC curve and F1, precision, recall
Khosla et al. [16]	2121	MobileNet, InceptionV3, and InceptionRestNetV2	Facial Images	Accuracy Score
Mishra et al. [20]	2021	Random Forest and Decision Tree	ABIDE-I	Accuracy, ROC curve and F1, precision, recall
Hosseini et al. [11]	2022	MobileNet	Facial Images	Accuracy score

**3. Proposed Methodology.** We divided the methodology portion in to subsection to elaborate the purpose of each section.

**3.1. Dataset Description.** Dataset Description. In this paper, we use the open-source dataset available on Kaggle [1]. The dataset is divided in to three groups for training, validation, and evaluation with the ratio of 86.38%, 10.22%, and 3.41%, respectively. Training set is used to train the model, validation to check the model, whilst on the test model we validate the effectiveness of proposed technique. The training set has 2536 face images, validation set has 300 and the test set has 100 images. While non-autistic face photographs were haphazardly gathered from the internet, web sources with issues with autism were used to obtain the youngsters with autistic faces. This dataset includes 2D RGB photos of kids between the ages of 2 and 14, with the majority falling between those two and eight. The dataset showed an around 3:1 male to female ratio, compared to a nearly 1:1 ratio for the autistic class and normal control class (Fig. 3.1).

**3.2. Pre-processing.** The duplicate photographs were removed from the dataset, and the images were cropped to just display the facial portion. We normalise the dataset between 0 and 1 by using standard PyTorch function transforms.

**3.3. Feature Extraction and Classification.** Thanks to the deep learning algorithms that made possible automatic feature extraction. The effort of extracting features from images is challenging, but with the advent of architecture based on convolutional neural networks, this task has become much easier. To categorize photos of autism, we are using convolutional neural networks, which apply deep learning methods for classification problems. The following layer types make up their layered structure: -

- Max Pooling Layer
- Sub-sampling Layer of the Convolution Layer
- Integrated Connection Layer

We used EfficientNet [31] in this study to detect autism because transfer learning leverages pretrained models from the ImageNet dataset rather than training from scratch. EfficientNet, one of the most efficient CNN models, exhibits exceptional accuracy on both ImageNet and common picture classification tasks using transfer learning while utilising the least number of FLOPS for inference. The EfficientNet B0 model, which is the foundational EfficientNet model, is used in this study with an input image size of (224 X 224 X 3). Dropout layers and batch normalisation can be added to the EfficientNet model during fine-tuning to help with overfitting issues. However, in our case, we just change the last layer of the classifier to make the binary classification task. The overall summary of the proposed model is shown in Fig. 3.2.

**3.4. Method of Instruction.** On the training dataset, we used batch-trained to the model, and the test dataset served as the model's evaluation. Utilising the fit-generator method, we created a special dataset function for this. The model is then developed and assembled using the training data and predefined hyper-parameters.

**3.5. Rate of Learning.** To decay, we employ the learning rate. Modern neural networks are trained using the learning rate decay technique, and if there is no change in loss values, the learning rate value is changed.



Fig. 3.1: Face Images

Table 3.1: Summary of Hyper\_parameter

Hyper-parameter	Value
Batch_size	24
No of epochs	300
Learning rate	0.001
Optimizer	Adam
Dropout value	0.5

The neural network is first trained with a high learning rate, which is subsequently decayed until local minima are discovered. It has been demonstrated to help with both generalisation and optimization.

**3.6. Optimization.** One of the key ideas in deep learning is cost function optimization. The gradient descent algorithm is the most popular one. The little dataset in our situation, however, makes it quite slow. We employ a variant of this method called Adam to assist our model in learning far more quickly.

**3.7. Loss Function.** The suggested study uses the cross-entropy loss function and relies on the binary categorization of images. Finding the differences between two probability distributions is its main goal. We employed the sigmoid function for the activation function.

EfficientNet (EfficientNet)	[32, 3, 224, 224]	[32, 2]	--	Par
Sequential (features)	[32, 3, 224, 224]	[32, 1280, 7, 7]	--	Fal
se				
↳Conv2dNormActivation (0)	[32, 3, 224, 224]	[32, 32, 112, 112]	--	Fal
↳Conv2d (0)	[32, 3, 224, 224]	[32, 32, 112, 112]	(864)	Fal
↳BatchNorm2d (1)	[32, 32, 112, 112]	[32, 32, 112, 112]	(64)	Fal
↳SiLU (2)	[32, 32, 112, 112]	[32, 32, 112, 112]	--	--
↳Sequential (1)	[32, 32, 112, 112]	[32, 16, 112, 112]	--	Fal
↳MBConv (0)	[32, 32, 112, 112]	[32, 16, 112, 112]	(1,448)	Fal
↳Sequential (2)	[32, 16, 112, 112]	[32, 24, 56, 56]	--	Fal
↳MBConv (0)	[32, 16, 112, 112]	[32, 24, 56, 56]	(6,004)	Fal
↳MBConv (1)	[32, 24, 56, 56]	[32, 24, 56, 56]	(10,710)	Fal
↳Sequential (3)	[32, 24, 56, 56]	[32, 40, 28, 28]	--	Fal
↳MBConv (0)	[32, 24, 56, 56]	[32, 40, 28, 28]	(15,350)	Fal
↳MBConv (1)	[32, 40, 28, 28]	[32, 40, 28, 28]	(31,290)	Fal
↳Sequential (4)	[32, 40, 28, 28]	[32, 80, 14, 14]	--	Fal
↳MBConv (0)	[32, 40, 28, 28]	[32, 80, 14, 14]	(37,130)	Fal
↳MBConv (1)	[32, 80, 14, 14]	[32, 80, 14, 14]	(102,900)	Fal
↳MBConv (2)	[32, 80, 14, 14]	[32, 80, 14, 14]	(102,900)	Fal
↳Sequential (5)	[32, 80, 14, 14]	[32, 112, 14, 14]	--	Fal
↳MBConv (0)	[32, 80, 14, 14]	[32, 112, 14, 14]	(126,004)	Fal
↳MBConv (1)	[32, 112, 14, 14]	[32, 112, 14, 14]	(208,572)	Fal
↳MBConv (2)	[32, 112, 14, 14]	[32, 112, 14, 14]	(208,572)	Fal
↳Sequential (6)	[32, 112, 14, 14]	[32, 192, 7, 7]	--	Fal
↳MBConv (0)	[32, 112, 14, 14]	[32, 192, 7, 7]	(262,492)	Fal
↳MBConv (1)	[32, 192, 7, 7]	[32, 192, 7, 7]	(587,952)	Fal
↳MBConv (2)	[32, 192, 7, 7]	[32, 192, 7, 7]	(587,952)	Fal
↳MBConv (3)	[32, 192, 7, 7]	[32, 192, 7, 7]	(587,952)	Fal
↳Sequential (7)	[32, 192, 7, 7]	[32, 320, 7, 7]	--	Fal
↳MBConv (0)	[32, 192, 7, 7]	[32, 320, 7, 7]	(717,232)	Fal
↳Conv2dNormActivation (8)	[32, 320, 7, 7]	[32, 1280, 7, 7]	--	Fal
↳Conv2d (0)	[32, 320, 7, 7]	[32, 1280, 7, 7]	(409,600)	Fal
↳BatchNorm2d (1)	[32, 1280, 7, 7]	[32, 1280, 7, 7]	(2,560)	Fal
↳SiLU (2)	[32, 1280, 7, 7]	[32, 1280, 7, 7]	--	--
↳AdaptiveAvgPool2d (avgpool)	[32, 1280, 7, 7]	[32, 1280, 1, 1]	--	--
↳Sequential (classifier)	[32, 1280]	[32, 2]	--	Tru
↳Dropout (0)	[32, 1280]	[32, 1280]	--	--
↳Linear (1)	[32, 1280]	[32, 2]	2,562	Tru

Fig. 3.2: Summary of EfficientNetB0

**4. Results and Discussion.** We utilized the following evaluation metrics to assess the performance of the suggested classifiers.

1. **Accuracy:** The degree to which a classifier can correctly predict the class for a given input is known as classification accuracy. It is described as the proportion of accurate predictions out of all possible predictions that the classifier made.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (4.1)$$

2. **Precision:** A classifier's precision, which is measured using a metric, is how well it predicts the positive class. Its definition is the proportion of the classifier's true positive predictions among all its positive

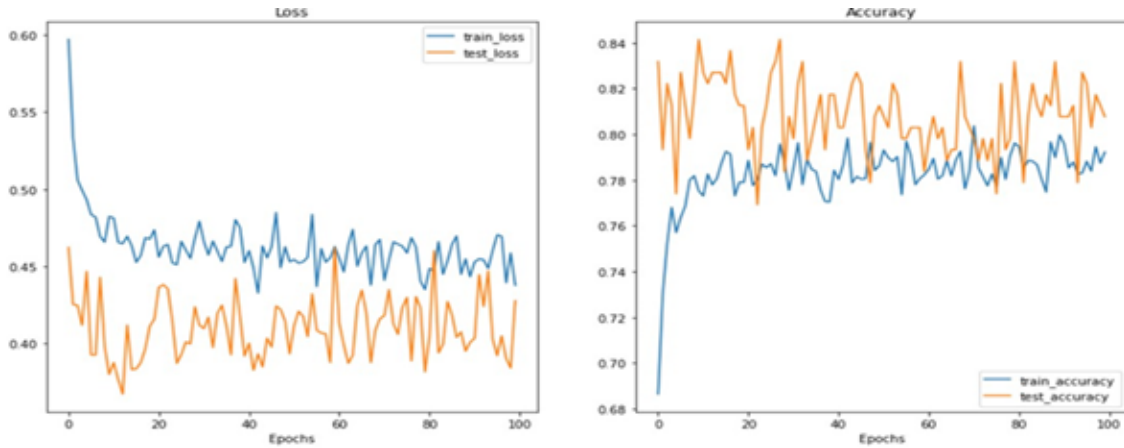


Fig. 4.1: Loss and Accuracy Curve

predictions.

$$Precision = \frac{TP}{TP + FP} \quad (4.2)$$

3. **Recall:** The metric of recall assesses a classifier's ability to correctly identify each occurrence of the positive class. It is described as the number of true positive predictions made by the classifier out of all the positive cases in the dataset.

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

4. **F1 Measure:** F1 measure is harmonic mean of recall and precision.

$$F1 - measure = 2 \frac{Recall \times Precision}{Recall + Precision} \quad (4.4)$$

We demonstrate the effectiveness of the suggested technique in terms of accuracy, precision, memory, F1 measure, confusion matrix, and AUC score. The loss and accuracy curve from Fig. 4.1 indicates initially accuracy was high and gradually model shows smooth accuracy, and the maximum predicted value of accuracy is 0.85%. While in case of loss, training loss was initially significantly higher compared to test loss. Training loss gradually decreases, and the problem of overfitting is resolved. We also examine the batch accuracy and batch loss data. We identified the model's bias toward non-autistic class cases in terms of accuracy. We also evaluate the effectiveness of the suggested method in terms of recall and precision and F1-score. The bar graph of precision shows the proposed module is biased toward the non-autistic samples and shows more instances in predicting the non-autistic. While the recall is more biased toward the autistic samples and predicts more samples of autistic class. The F1 is average in both precision and recall. The overall description of classification report is shown in figure 4.2. We also display the confusion matrix on the test dataset, which displays instances that were properly and incorrectly predicted in figure 4.3. We also make comparison with one study to evaluate the effective of proposed technique

**5. Conclusion.** We can wrap up by mentioning this paper's three main contributions. (i) To address the issue of transfer learning, we presented an approach by using last three layers of EfficientNetB0 for detecting autism disorder as not having enough training data. The method is based on training the deep learning model with a balanced approach on a little amount of dataset (ii) Our designed approach is simple, efficient, and beats the baseline transfer learning techniques, (iii) we proposed techniques that showed outstanding results

Table 4.1: Comparative Analysis

Study	Model	Accuracy Score
Jahanara et al. [12]	VGG19	0.84
Our	EfficientNetB0	0.85

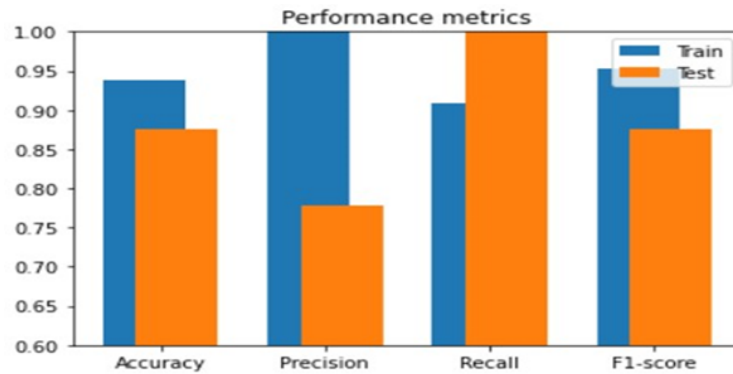


Fig. 4.2: Classification Report

in terms of accuracy, with accuracy scores of 85% , precision , recall and F1-measure. This research shows that predictive analysis based on transfer learning using EfficientNetB0 from autism images is highly efficient and provides a simple path for automatic detection. In future, we are interested to explore the generative and transformer-based approaches for prediction of autism.

## REFERENCES

- [1] <https://drive.google.com/drive/folders/1xqu0plul0m3tilxqntano12d68pemb8a>.
- [2] <https://www.kaggle.com/c/challenges-in-representation-learning-facial-expression-recognition-challenge>.
- [3] Z. A. AHMED, T. H. ALDHYANI, M. E. JADHAV, M. Y. ALZHRANI, M. E. ALZHRANI, M. M. ALTHOBAITI, F. ALASSERY, A. ALSHAFLUT, N. M. ALZHRANI, AND A. M. AL-MADANI, *Facial features detection system to identify children with autism spectrum disorder: deep learning models*, Computational and Mathematical Methods in Medicine, 2022 (2022).
- [4] F. W. ALSAADA AND M. S. ALZHRANI, *Classification and detection of autism spectrum disorder based on deep learning algorithms*, Computational Intelligence and Neuroscience, 2022 (2022).
- [5] S. BEBORTTA, S. S. TRIPATHY, S. BASHEER, AND C. L. CHOWDHARY, *Deepmist: Towards deep learning assisted mist computing framework for managing healthcare big data*, IEEE Access, (2023).
- [6] M. BOTHA, B. DIBB, AND D. M. FROST, "autism is me": an investigation of how autistic individuals make sense of autism and stigma, *Disability & Society*, 37 (2022), pp. 427–453.
- [7] K. CAMPBELL, K. L. CARPENTER, S. ESPINOSA, J. HASHEMI, Q. QIU, M. TEPPER, R. CALDERBANK, G. SAPIRO, H. L. EGGER, J. P. BAKER, ET AL., *Use of a digital modified checklist for autism in toddlers—revised with follow-up to improve quality of screening for autism*, *The Journal of Pediatrics*, 183 (2017), pp. 133–139.
- [8] F. CHIAROTTI AND A. VENEROSI, *Epidemiology of autism spectrum disorders: a review of worldwide prevalence estimates since 2014*, *Brain sciences*, 10 (2020), p. 274.
- [9] T. GHOSH, M. H. AL BANNA, M. S. RAHMAN, M. S. KAISER, M. MAHMUD, A. S. HOSEN, AND G. H. CHO, *Artificial intelligence and internet of things in screening and management of autism spectrum disorder*, *Sustainable Cities and Society*, 74 (2021), p. 103189.
- [10] M. I. U. HAQUE AND D. VALLES, *A facial expression recognition approach using dcnn for autistic children to identify emotions*, in 2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), IEEE, 2018, pp. 546–551.
- [11] M.-P. HOSSEINI, M. BEARY, A. HADSELL, R. MESSERSMITH, AND H. SOLTANIAN-ZADEH, *Deep learning for autism diagnosis and facial analysis in children*, *Frontiers in Computational Neuroscience*, (2022), p. 119.
- [12] S. JAHANARA AND S. PADMANABHAN, *Detecting autism from facial image*, (2021).
- [13] Y. JIN, C.-Y. WEE, F. SHI, K.-H. THUNG, D. NI, P.-T. YAP, AND D. SHEN, *Identification of infants at high-risk for autism spectrum disorder using multiparameter multiscale white matter connectivity networks*, *Human brain mapping*, 36 (2015),

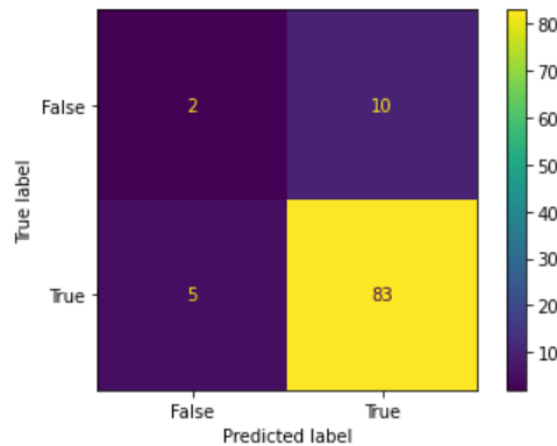


Fig. 4.3: Confusion matrix

- pp. 4880–4896.
- [14] Y. F. KHAN, B. KAUSHIK, C. L. CHOWDHARY, AND G. SRIVASTAVA, *Ensemble model for diagnostic classification of alzheimer's disease based on brain anatomical magnetic resonance imaging*, *Diagnostics*, 12 (2022), p. 3193.
- [15] M. KHODATARS, A. SHOEIBI, D. SADEGHI, N. GHAASEMI, M. JAFARI, P. MORIDIAN, A. KHADEM, R. ALIZADEHSANI, A. ZARE, Y. KONG, ET AL., *Deep learning for neuroimaging-based diagnosis and rehabilitation of autism spectrum disorder: a review*, *Computers in Biology and Medicine*, 139 (2021), p. 104949.
- [16] Y. KHOSLA, P. RAMACHANDRA, AND N. CHAITRA, *Detection of autistic individuals using facial images and deep learning*, in 2021 IEEE International Conference on Computation System and Information Technology for Sustainable Solutions (CSITSS), IEEE, 2021, pp. 1–5.
- [17] B. LI, S. MEHTA, D. ANEJA, C. FOSTER, P. VENTOLA, F. SHIC, AND L. SHAPIRO, *A facial affect analysis system for autism spectrum disorder*, in 2019 IEEE International Conference on Image Processing (ICIP), IEEE, 2019, pp. 4549–4553.
- [18] R. MAROTTA, M. C. RISOLEO, G. MESSINA, L. PARISI, M. CAROTENUTO, L. VETRI, AND M. ROCCELLA, *The neurochemistry of autism*, *Brain sciences*, 10 (2020), p. 163.
- [19] P. MCCARTY AND R. E. FRYE, *Early detection and diagnosis of autism spectrum disorder: why is it so difficult?*, in *Seminars in Pediatric Neurology*, vol. 35, Elsevier, 2020, p. 100831.
- [20] M. MISHRA AND U. C. PATI, *Autism spectrum disorder detection using surface morphometric feature of smri in machine learning*, in 2021 8th international conference on smart computing and communications (icscc), IEEE, 2021, pp. 17–20.
- [21] L. MOTTRON AND D. BZDOK, *Autism spectrum heterogeneity: fact or artifact?*, *Molecular psychiatry*, 25 (2020), pp. 3178–3185.
- [22] K. MUJEEB RAHMAN AND M. M. SUBASHINI, *Identification of autism in children using static facial features and deep neural networks*, *Brain Sciences*, 12 (2022), p. 94.
- [23] I. J. OOSTERLING, M. WENSING, S. H. SWINKELS, R. J. VAN DER GAAG, J. C. VISSER, T. WOUDEBERG, R. MINDERAA, M.-P. STEENHUIS, AND J. K. BUITELAAR, *Advancing early detection of autism spectrum disorder by applying an integrated two-stage screening approach*, *Journal of Child Psychology and Psychiatry*, 51 (2010), pp. 250–258.
- [24] S. PADINJAPPURATHU GOPALAN, C. L. CHOWDHARY, C. IWENDI, M. A. FARID, AND L. K. RAMASAMY, *An efficient and privacy-preserving scheme for disease prediction in modern healthcare systems*, *Sensors*, 22 (2022), p. 5574.
- [25] S. PARISOT, S. I. KTEA, E. FERRANTE, M. LEE, R. GUERRERO, B. GLOCKER, AND D. RUECKERT, *Disease prediction using graph convolutional networks: application to autism spectrum disorder and alzheimer's disease*, *Medical image analysis*, 48 (2018), pp. 117–130.
- [26] P. K. C. PRASAD, Y. KHARE, K. DADI, P. VINOD, AND B. R. SURAMPUDI, *Deep learning approach for classification and interpretation of autism spectrum disorder*, in 2022 International Joint Conference on Neural Networks (IJCNN), IEEE, 2022, pp. 1–8.
- [27] D. A. ROSSIGNOL AND R. E. FRYE, *Melatonin in autism spectrum disorders: a systematic review and meta-analysis*, *Developmental Medicine & Child Neurology*, 53 (2011), pp. 783–792.
- [28] S. R. SHAHAMIRI AND F. THABTAH, *Autism ai: a new autism screening system based on artificial intelligence*, *Cognitive Computation*, 12 (2020), pp. 766–777.
- [29] B. K. SWAIN, M. Z. KHAN, C. L. CHOWDHARY, AND A. ALSAEDI, *Src: Superior robustness of covid-19 detection from noisy cough data using gfcc*.
- [30] K. TAIT, F. FUNG, A. HU, N. SWELLER, AND W. WANG, *Understanding hong kong chinese families' experiences of an autism/asd diagnosis*, *Journal of Autism and Developmental Disorders*, 46 (2016), pp. 1164–1183.
- [31] M. TAN AND Q. LE, *Efficientnet: Rethinking model scaling for convolutional neural networks*, in International conference on

machine learning, PMLR, 2019, pp. 6105–6114.

*Edited by:* Chiranji Lal Chowdhary

*Special Issue on:* Scalable Machine Learning for Health Care: Innovations and Applications

*Received:* Apr 12, 2023

*Accepted:* Jun 6, 2023





## PRIVACY AND SECURITY ENHANCEMENT OF SMART CITIES USING HYBRID DEEP LEARNING-ENABLED BLOCKCHAIN

**Abstract.** The emergence of the Internet of Things (IoT) accelerated the implementation of various smart city applications and initiatives. The rapid adoption of IoT-powered smart cities is faced by a number of security and privacy challenges that hindered their application in areas such as critical infrastructure. One of the most crucial elements of any smart city is safety. Without the right safeguards, bad actors can quickly exploit weak systems to access networks or sensitive data. Security issues are a big worry for smart cities in addition to safety issues. Smart cities become easy targets for attackers attempting to steal data or disrupt services if they are not adequately protected against cyberthreats like malware or distributed denial-of-service (DDoS) attacks. Therefore, in order to safeguard their systems from potential threats, businesses must employ strong security protocols including encryption, authentication, and access control measures. In order to ensure that their network traffic remains secure, organizations should implement powerful network firewalls and intrusion detection systems (IDS). This article proposes a blockchain-supported hybrid Convolutional Neural Network (CNN) with Kernel Principal Component Analysis (KPCA) to provide privacy and security for smart city users and systems. Blockchain is used to provide trust, and CNN enabled with KPCA is used for classifying threats. The proposed solution comprises three steps, preprocessing, feature selection, and classification. The standard features of the datasets used are converted to a numeric format during the preprocessing stage, and the result is sent to KPCA for feature extraction. Feature extraction reduces the dimensionality of relevant features before it passes the resulting dataset to the CNN to classify and detect malicious activities. Two prominent datasets namely ToN-IoT and BoT-IoT were used to measure the performance of this anticipated method compared to its best rivals in the literature. Experimental evaluation results show an improved performance in terms of threat prediction accuracy, and hence, increased security, privacy, and maintainability of IoT-enabled smart cities.

**Key words:** Blockchain, Deep Learning, Convolutional neural network, Principal Component Analysis, Privacy and security, Intrusion detection, Internet of Things, Sensor technology

**AMS subject classifications.** 68M12, 68T05

**1. Introduction.** The increase in urbanization in recent years has demanded the economic, ecological, and social expansion of metropolises to improve one's quality of life meaningfully. This has brought about the introduction of the smart city concept to bring about development in urban cities. According to the United Nations assessment, by 2050, more than 70% of the world's population would be living in cities [1]. In various fields, like medicine, transport, teaching, budget, working conditions, and living environments, the development procedure has significantly enhanced people's living principles [2]. Nonetheless, the high density of the population poses significant hurdles in terms of distributing available resources using brand-new technologies. The swift growth of the urban population thus has an impact on eco-friendly reserve limitations like air smog, transportation jamming, garbage dumping, and greenhouse gas emissions. For cities' long-term development, all of these concerns and challenges necessitate innovative answers, bringing the notion of "Smart City", SC, into play [3, 4, 5].

Thus, the integration of the Internet of Things (IoT) and other tools have been used to resolve urban issues and problems [6, 7]. IoT-enabled applications are critical for making the best use of existing resources and technologies in the development of SCs. However, it has a single-point-of-failure problem that causes safety, confidentiality, latency, and reliability difficulties. Using IoT intelligent services provides chances to address the aforementioned difficulties and create a high-quality living. Due to qualities like transparency, trust-free, decentralization, and immutability. While providing high-quality services, blockchain tools can be utilized to address the aforementioned danger and privacy complications. Within today's (4G/5G) and next-generation (6G) wireless systems, IoT is likely to become the dominant trend in Internet and service-centric computing, which will play a large role in the next generation of sustainable smart cities.

Appropriate with larger collections of ecological, commercial, and community data conveyed via IoT, con-

nectivity among resources in IoT-enabled SCs may be built and refined [8, 9, 10]. As a result, ecologically sustainable living and governance become a reality using the smart cities paradigm. One of the primary aiding tools for SC applications is the IoT which makes use of Internet technologies to connect smart devices and items [11, 13, 12]. In such an IoT model, data is acquired through a variety of physical devices, dispersed through wireless systems, and then handled in real time. Actuators are controlled using the data collected and processed [14]. As a result, in the comprehensive growth of IoT connectivity in smart cities, safety, secrecy, reliance, scalability, and centralization are all critical challenges that must be addressed [15, 16]. IoT systems, on the other hand, are highly dispersed and diverse, making them different from traditional systems.

The sole characteristics of IoT, computing power, memory volume, network bandwidth, and battery lifespan have made IoT security, privacy, and trust in SC design harder to maintain [17]. However, the interconnectedness of numerous IoT devices in smart systems creates a wide range of possible attacks aimed at IoT devices in SCs. The two main types of attacks in a smart city are cyber and physical. Physical attacks are launched using devices in the network where attackers can modify or temper the sensors and devices since the invaders are nearer to the devices [17, 18]. There are various attacks like Sleep Denial Attacks, malicious code injection, Permanent Denial of Service, radio frequency congestion, and false node injection in the smart city [19, 20]. In cyber-attacks, the attackers can inject malware or malicious software to gain illegal entrance to the structures of the network schemes [21].

Attacks like Ransomware, Denial-of-Service (DoS), Man-In-The-Middle, and Distributed Denial-of-Service (DDoS) are some examples of cyber-attacks targeting a smart city [22, 23]. Such attacks are becoming more common at an alarming rate, posing a threat to data security, reliability, and accessibility. Furthermore, the expanding privacy concerns include the vulnerability of sensitive data through inference and data-harming attacks, which can affect smart cities. For instance, attackers attempt to adjust IoT devices in data and inference poisoning attacks by adding made-up data measurements [24]. As a result, regular communications between smart things are hampered, and the power of smart devices is wasted [25]. Additionally, it is possible that such attacks could take a detrimental effect on the performance of data analysis systems that use machine learning (ML), e.g., intrusion detection systems (IDS).

As a result, privacy is critical in the operation of SCs and associated networks. The architecture encompassing security, privacy, and trust can be deemed a sufficient method to address the existing challenges. Hence, IoT-based SCs and their networks require a reliable and effective security system. To improve the eminence of the data in IoT-based systems, the use of redundant sensors has been generally adopted. On the other hand, the use of various sensors, reliable, unreliable, or corrupted, can provide many readings or observations of the same thing [26, 27]. Hence, it is expected that it is needed to deal with and find a permanent solution to those vulnerable and unreliable IoT nodes, thus providing reality and trust from unreliable results [28]. In the case of addressing such security issues, solutions could result in high computation costs of ML-based IDS models [29], privacy violations of Cloud-IoT applications [30], and/or a high false alarm rate of IDS models [31]. Therefore, it has become important to develop holistic smart cities that will maintain trustworthiness among IoT nodes.

Therefore, to overcome the aforementioned issues, this paper proposes a blockchain-enabled hybrid convolutional neural network CNN with Kernel Principal Component Analysis (KPCA) for the confidentiality security and privacy of SCs. Blockchain technology is a viable method to support cloud computing by making a distributed cloud for IoT-enabled smart cities. This will completely identify and measure the cloud infrastructure by managing the cloud and holding it responsible for its actions, and allow consumers of the IoT-enabled stage to confirm that the environment is working perfectly and in real-time.

This was achieved using a two-level security-preserving method. Blockchain technology is used at the first level for data authentication and for the prevention of captured data from poisoning attacks. The Kernel Principal Component Analysis (KPCA) method was used in the second phase for the choice of important characteristics and attributes from the datasets are transformed into an encoded format for preventing assumption attacks based on the deep learning technique used and improving the overall effectiveness of the proposed system. The obtained reduced features are used as an input to the CNN for the classification of normal and attack, and the blockchain is used for the authentication and privacy of the IIoT-based systems. Two prominent datasets namely ToN-IoT [32] and BoT-IoT [33, 34] were used to evaluate the security mechanism since both datasets have various IoT-based attacks like DoS, Ransomware, normal vectors, MITM, and Theft as mentioned

earlier, and are publicly available [35].

The following are the study contributions:

- (i) The paper proposed a privacy and trustworthy framework using blockchain technology to provide dependability within the device layer of IoT architecture, and KPCA was used for feature reduction with blockchain-based enabled Proof of Work (ePoW) for the protection of IoT-based systems from inference and poisoning attacks.
- (ii) CNN was used for the classification of data for the detection of suspicious activities within the smart city networks.
- (iii) The paper used blockchain-enabled to prevent the cloud database from the problems of redundancy in IIoT-based data using an on-chain technique called CloudBlock-EdgeBlock architecture to deploy the proposed system since it supports verifiability services, and traceability within IoT nodes in a smart city, timestamp records are generated for each transaction. The CloudBlock-EdgeBlock architecture was used to enhance blockchain-enabled on-chain devices in order to support indisputability, verifiability, and traceability by generating timestamp records for each transaction performed in the smart city within IIoT-based nodes. This was done to address the problems of Cloud-fog redundancy in IIoT data.
- (iv) The proposed system efficiency is measured using F1-score, precision detection rate, and accuracy. Two publicly datasets BoT-IoT and BoT-IoT are used for the experimental performance and the results were compared with most recently related systems using non-blockchain and blockchain systems.

The remaining part of the paper is organized as follows: section 2 discussed the Blockchain-enabled AI in IoT-based Smart Cities. Section 3 presented the Artificial Intelligence enabled Intrusion Detection for IoT in Smart Cities. Section 4 reviewed related work in the areas of Blockchain, and applications of deep learning in smart cities. Section 5 presents the methodology employed in this study. Section 6 presented the experimental results and also discusses the performance analysis of the study. Finally, section 7 concluded the study with future and open research in the area.

**2. The Blockchain-enabled with Artificial Intelligence in the Internet of Things for Smart Cities.** In order to provide improved services to its citizens while ensuring effective and optimal use of available resources, for the combination and administration of physical, social, and corporate infrastructures, a smart city requires information technology [36]. The IoT is a concept that enables humans and devices to communicate via the internet. Smart houses, intelligent automobiles, smart industries, and smart transportation are among the equipment capable of intercommunication [37]. The IoT offers a variety of solutions for many fields, as depicted in fig.2.1, to help them maximize their output more effectively and efficiently. Despite its many benefits, IoT is plagued by challenges including data protection, centralized, data analytics, connectivity, and more technology limitations. In 2015, over 800,000 user devices were found to have been infected as a result of spam emails and distributed phishing [39].

The author in [36] claimed that as the quantity of smart devices grows, so does the amount of data created. As a result, big data analytics is an important consideration for any IoT application. Various studies contribute different solutions for IoT applications using technologies including Artificial Intelligence (AI) and Deep Learning (DL) to address this issue of data analytics [40]. The DL approach is utilized to evaluate massive volumes of data in order to provide information for outcome, forecasting, and categorization procedures. The DL analytics tool can extract and scale features from a large volume of data from IoT applications [41].

Blockchain technology creates a network that is decentralized, distributed, and secure. In blockchain tools, each node is linked in a decentralized peer-to-peer network, where each transaction is immediately transactions are shared without the use of timestamps and documented with timestamps influence. Agriculture, healthcare, security, and finance are all areas where the blockchain method might be beneficial. Through cryptographic hashing, the data provided in blocks is additionally involved and safeguarded in chains with digital signs. Because each block is connected to the one before it, hackers will be unable to hack transactions by injecting harmful data into the system. Digital signatures, validation, smart contracts, decentralization, and immutable explainable AI have all been handled by combining the blockchain approach with artificial intelligence for IoT frameworks. With the development of smart IoT devices and their interconnections, massive amounts of data are now being generated in a consolidated format.

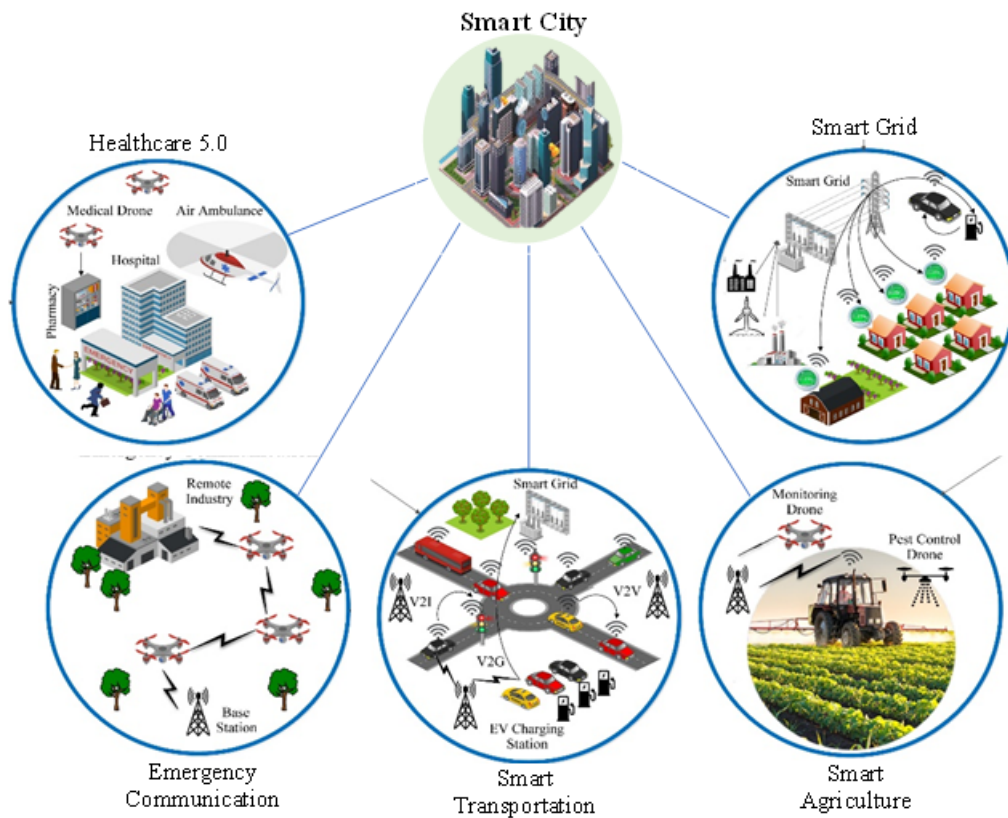


Fig. 2.1: Application-specific view of Smart Cities

As a result, technological advancements frequently generate difficulties such as space, security, and privacy. A decentralized database system is being constructed with the combination of blockchain and AI for IoT to address these challenges [42]. While sharing the transaction with anyone else on the network, it should be safe, digitally signed, verified, and transparent. A secure transaction model like this can be used in a variety of applications, including healthcare, smart homes, agriculture, military, industrial, smart transportation, and many more. [43]. For strengthening network security, blockchain equipment implements the concept of smart contracts, which are then kept in a digital ledger [44]. AI is used in a variety of advanced technology domains; decentralized AI, blockchain (BC), IoT intelligence, machine automation, and so on are only a few examples. The combination of AI and the IoT offers advantages in terms of data collection and processing [45]. AI, blockchain, and the IoT have attracted interest from numerous academics in recent years as smart and digital technologies have evolved, and they have risen to prominence as the most widely used technology, generating innovative ideas in a variety of research domains [39]. Fig.2.2 illustrates the basic confluence of blockchain and AI for IoT applications.

The concept of SCs is slowly but steadily becoming a reality as various countries throughout the world adopt and create their own smart city models. At the heart of the smart city are the devices and actuators built into smart gadgets that detect the location and help people make better decisions. The microchips in these devices have been designed to make decisions on their own based on the data collected from the sensors. This entails combining various technologies, like AI, protocols, IoT, wireless sensor networks (WSN), and so on. The AI, IoT, ML, DL, and the terms "cognitive computing" and "big data analytics" have been used interchangeably all played a role in making this goal a reality [46, 47].

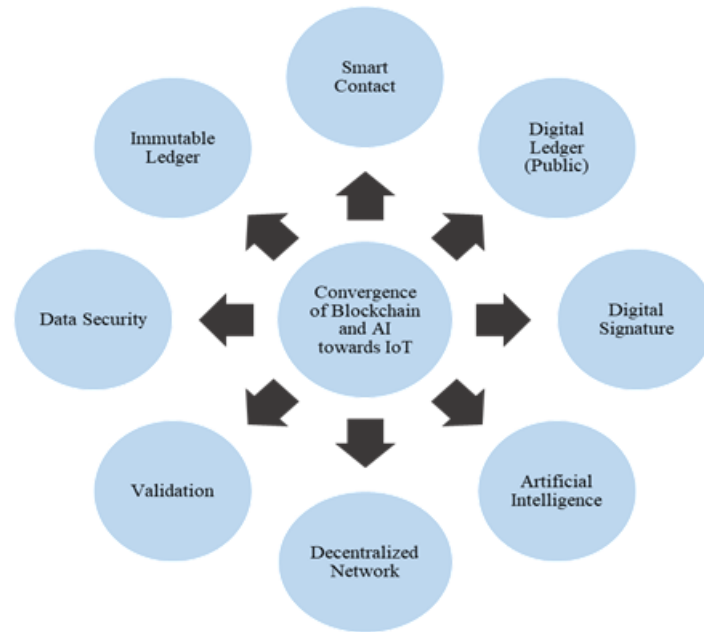


Fig. 2.2: The basic convergence of blockchain and AI for IoT applications.

One such promising initiative that has been implemented globally with the goal of making residents' life more convenient and inclusive is a smart city [48, 49]. The concept is to employ current technology to transform each object of a traditional city into an independent entity that can run on its own without any external assistance. All daily activities, such as authority, strategies, services, and responses, are computerized, and operators can access them via smart devices from anywhere on the globe. By employing environmentally safe and cost-effective strategies, automation has aided in the reduction of environmental dangers.

**3. Artificial Intelligence enabled Intrusion Detection for IoT in Smart City.** The increased use of the internet in recent years globally has made it a favorable environment for nefarious activity. Malicious IoT-based systems for SCs are an example of these activities, and they are one of the greatest severe pressures to internet users and smart cities. This rapid development has dramatically increased the number of urbanities that moved to the internet to create a smart city. The fast-growing in these areas expanded online crimes that run using malicious network nodes. Detection and analysis of these networks became one of the major problems of online space [40]. Many professional users that are conscious of malicious activities have also been victims of such attacks.

Typically, malicious pages are online attacks trying to steal users' sensitive information by making a malicious network node and leading users to that, so users may think they interact with a legitimate IoT platform while using a malicious one. Presently, there are many types of malicious activities in IoT-based systems, and they can appear to be any normal network, such as phishing websites and malware-hosting or malware propagation [50]. The SC is one of the core areas of the IoT field, burgeoning and integrated into daily human events. A smart city is defined as an internet-enabled city for the remote control, monitoring, and management of city appliances via a designated device like a smartphone, tablet, or laptop. A smart city environment, otherwise called an intelligent city, city automation, or domotics, offers users comfortability, convenience, city safety, security, and energy efficiency [51].

According to authors in [52], the smart city incorporates communication networks for connecting major city appliances and services for remote control, monitoring, and access of the residence from within or outside the premises. The major benefits associated with smart city automation systems are healthy living with its deployment in healthcare systems, reduction in energy consumption, and home safety and security. Some

countries support the full deployment of smart city automation systems for their residents by putting laws, rules, and subsidies to facilitate its use and encourage residents to adopt a decrement in the rate of energy demand [53]. The increasing interest in research about the city automation system has proven that it is an area that needs exploration to enhance its functionality and accessibility. However, some challenges are still mitigating against the extensive use of smart city automation systems: cost, data storage security, and secured communication channels. This has led to various research works to give an in-depth discussion of smart city automation and find solutions to the challenges of safety and confidentiality in the smart city. The IoT-enabled smart city has been subjected to a variety of cyber-attacks, putting its ability to provide flawless operations to metropolitan areas in jeopardy. Users and city automation suffer financial and reputational harm as a result of such threats, as well as the stealing of confidential documents.

As a result, several Intrusion Detection Systems (IDSs) have been established to combat and defend IoT-enabled smart city systems. However, gathering data that can be used in the advancement of a sophisticated IDS is a tough mission, and there are significant problems in identifying prevailing and novel assaults [20, 40]. As the quantity of IoT-based smart city devices and applications grows, protecting important public infrastructure is becoming a more pressing issue in every city [54]. Malware that takes advantage of zero-day vulnerabilities is one of the most common threats in IoT networks using several ways, the offenders infect vulnerable machines to track and change their behavior [20]. These malicious behaviors demonstrated that traditional cyber-threat methods are no longer adequate for protecting crucial infrastructure such methods are weak and not able to recognize the threats in real time. The NIDS is critical in identifying and responding to any online threats as a cybersecurity mechanism.

The IoT-enabled smart city has evolved into a critical component of today's data and information transfer machinery, prompting the need for global network security [56]. NIDS are frequently used to detect system traffic to protect workstation schemes against numerous grid invasions. Intrusion, according to [40], is a framework for attempting to compromise an information system's security services. In reaction to the difficulties raised by these intrusive systems, researchers have been encouraged to develop novel IDSs. Several IDSs have been developed and enhanced in the past, but they remain susceptible to a diversity of occurrences. The potential of IDS to track and foresee malicious conduct and unknown assaults has sparked a surge of interest in anomaly detection research. Current machine learning-based irregularity detection algorithms, on the other hand, have a significant false alarm rate [57]. According to recent studies, feature selection is now at the heart of an added precise IDS [40, 59].

The feature selection strategy is utilized in most detection methods to choose the fitness values input characteristics for classification models, to improve overall finding performance and lower error rate in NIDS [60]. A classifier feature directions, in particular, are large, and not all of them relate to the groups to be classified, necessitating the adoption of a feature selection approach. The feature selection techniques, on the other hand, can be split into three categories [61]. The most prevalent feature selection technique relies on dataset characteristics without observing the classifier's effectiveness to select the most effective feature. The wrapper technique, on the other hand, is superior since it assesses the quality of the feature subclass using the classifier feedback, resulting in improved prediction performance. Similar to wrapper techniques, to improve the search efficiency of the learning algorithm, a classifier with an inherent process modeling function could be used in the integrated process. Several other Till now, IDS sections have been developed. The IDSs can be classed as rule-based or non-rule-based, misappropriation discovery, or various techniques, depending on the categorization algorithm used. IDSs can be characterized as real-time if they use permanent system tracking, or intermittent or passive if the traceability takes place only intermittently at set times or even offline utilizing data collected and analyzed over time.

The information received from the detection systems concerning the identified attacks is used to take countermeasures. The more precisely sorted and effective remedies selected, the less they hamper the device's or channel's regular functioning, and the more potent the assault is classed. Additionally, if the same type of attack is not identified, a counterattack may have more severe consequences than the attack itself. As a result, an IDS was created and demonstrated the work in each sort of attack. Furthermore, the system's false alarm rate is minimal, and its detection accuracy is good for both routine and irregular assaults, allowing for little processing to appropriately classify.

Because IDSs are utilized in the dependable and alert system of cyberattacks is crucial in industrial control systems that govern critical infrastructures [62], the latter attribute is essential. The feature extraction technique works well for designing and implementing authentic safety resolutions, as well as increasing IDS performance [63]. The necessity for improved correctness and a lower incorrect alarm rate in particular phenomenon discovery approaches developed the idea of data preprocessing and recognition as two reciprocal stages for IDS models [64, 65]. After eradicating superfluous features from the dataset and preserving a reduced feature set that may be utilized to produce a high-performance version, the preprocessing step employs the reduction attributes to remove the identification process. to use the base classifier to forecast attack kinds.

**4. Related Work.** The aim of this study is to integrate blockchain with AI for security, privacy, and intrusion detection in IoT-based enabled smart cities. Various works have been proposed to improve security, privacy, and trustworthiness within IoT-based nodes. Several studies have been published in the literature to promote trustworthiness among IoT nodes. In IoT-based systems, [66] suggested a trust architecture that integrated cross-layer permission protocol with Software Defined Network. A Trust Chain was presented by [67]. For allocating reliance and status scores between supply chain actors, this concept offers a three-layered trust management structure based on a consortium blockchain. To lower the number of computational resources used on the Internet of Vehicles [68] introduced consortia blockchain-based reserve distribution and a lightweight Proof-of-Reputation framework (IoV). Proof of reputation consensus procedure was proposed by [69].

The node's asset, consensus involvement, and transaction activity are used to generate the reputation score. A new block with the highest reputation is constructed based on the above score, and reputation-based voting is used to validate the new block. As a result, this method eliminates the need for miners. [70] developed a blockchain-based trust and status scheming system for safety and information critical domains in a distributed multi-agent framework. When agents engage with one another, a reputation score is calculated, and the score is kept on the blockchain to ensure that interactions between agents are trustworthy. Data pre-processing and Gaussian combination are employed for privacy, and ID uses the Kalman filter. A deep blockchain architecture for IoT was created by [71].

Bidirectional Long Short-Term Memory (BiLSTM) was utilized to create a privacy-preserving BC with smart contracts and irregularity discovery. The authors in [20] proposed an IIoT-DL-based ID framework with hybrid rule-based feature selection to train and verify information gathered from TCP/IP packets. A deep feedforward neural network model and a hybrid rule-based feature selection model were used in the training procedure. The results of the performance comparison show that the scheme outperforms other techniques with an accuracy and detection rate of 99.0%, and FPR of 1.0%, for the NSL-KDD dataset. In the case of the UNSW-NB15 dataset the accuracy, detection rate, and FPR of 98.9, 99.9, and 1.1% respectively. The researchers of suggested IDs in wireless connections in [72] and the Aegean AWID datasets were used to demonstrate the system's correctness. A PC, two workstations, one tablet, two cell phones, and an intelligent TV were used to collect the AWID dataset from a SOHO 802.11 wireless network protocol. However, the gathering only includes traces from the Media Access Control surface session and does not include data from IoT devices.

The researchers of [32] created a BoT-IoT dataset using a detailed simulation founded on an IoT network. DDoS, DoS, network check, and keylogging were among the acceptable and aggressive traffic retrieved, and data leakage is an instance of an attack that included both legal and unfriendly traffic. The internet traffic reported by the modeled IoT-based model utilizing the BoT-IoT dataset was more than 72 million. The study has presented a scaled-down version of the dataset with approximately 3.6 million entries for assessment purposes. A comparable dataset was used in [73] for ADS identification based on DoS assaults in an IoT network of sensors. The data acquired utilizing traditional and DoS assaults are SNMP/TCMP flooding, Ping of Death, and TCP SYN flooding, which simulated a smart home environment. However, because the dataset was not gathered using an IoT-based device, vulnerabilities such as XSS-Cross-site-site were not present in Spyware and scripting. The researchers of [74] suggested an ML-based approach for extracting malware pictures with a blend of local and global features.

The Mailing dataset was utilized to evaluate the performance of the proposed method, which includes 9339 samples from 25 ransomware families. After extracting features, the model showed a 99.21% accuracy rate and 98.40% precision classification utilizing 5288 samples from 8 ransomware families from the dataset. The researchers of [75] suggested a CNN model remove threats from a corpus of binary executables, and

their approach achieved 98.52 percent classification accuracy using the Mailing dataset of 9339 pieces from 25 ransomware. Aside from that, this pattern is utilized to select 10% of data arbitrarily for analysis of the dataset. In [76], the study proposes a ransomware detection technique based on CNN. On the same dataset, this model obtained a 98 percent accuracy rate. Within every cycle, a randomized mechanism is used to pick 10% of the samples to study the ransomware group in concern. The researchers advocated employing a Gaussian distribution for demographic initiation [77].

Moreover, to accomplish better discovery throughout each iteration, the Gaussian density function and the local-global best function were utilized in cooperation with the local technique. LGPLA-performance NN's was assessed to that of various current advanced approaches, including weight optimization utilizing Particle Swarm Optimization (PSO-NN) and BA-NN. The trial findings demonstrated that LGPLA-NN outperformed other variations in multi-class botnet security attacks, with an accuracy of 85.5 and 85.2 percent, accordingly. The researchers of [78] present a spyware detection technique based on ensemble learning. The foundation phase classification is performed by a stacking ensemble of fully integrated and one-dimensional CNNs, whereas the final classification is performed by an ML algorithm. The researchers compared and examined 15 ML classifications for a meta-learner in the research. In the assessment, five ML approaches were used: naive Bayes, decision tree, random forest, gradient boosting, and AdaBoosting.

Tests on the Windows Portable Executable malware dataset resulted in the following findings. The best results were obtained using an ensemble of seven neural network models plus the ExtraTrees classifier as a concluding classification model. [79] proposes a unique multilevel DL image classification method for intrusion detection systems. The network properties are transformed into four-channel images. The pictures are then used to train and assess the pre-trained DL model ResNet50. The suggested approach is tested against two publicly available benchmark datasets, UNSW-NB15 and BOUN Ddos. The suggested approach detects the general assault with 99.8 percent accuracy on the UNSW-NB15 dataset. On the BOUN DDoS dataset, the suggested model detects DDoS assaults with 99.7 percent accuracy and ordinary traffic with 99.7 percent accuracy.

The evaluation of existing related works has shown that there is a necessity for a better AI-based model to increase the accuracy and performance of IDS for IoT-based enabled smart cities. This motivates the use of a blockchain for security and privacy-preserving systems for the protection of IoT-based nodes in smart cities, KPCA for feature selection reduction, and CNN for the datasets classification, thus sensing attack trends within data as suspect vectors using depth coverage for data transmission. The proposed model will protect the captured data in SCs, and protect the process of IoT-based data sources within network traffic.

**5. Methods and Materials.** This section presents the proposed model by explaining in detail the system deployed to eliminate the problems of standalone architecture using cloud databases in IoT-based enabled smart cities. The system takes benefit of the mixture of blockchain and AI-based models for the security and privacy of the IoT-based smart city platform. The blockchain was used on the cloud side to protect the captured data that are stored on the cloud database using CloudBlock and EdgeBlock. The framework contains a two-level security-privacy-preservation method and intrusion detection using CNN techniques after employing KPCA for feature reduction for better accuracy. The sub-section discusses in detail the methodology.

The CloudBlock and EdgeBlock framework: the framework creates various data centers provided by several merchants. The CloudBlock is made up of numerous companies that offer various types of data facilities. The concept made use of three data facilities, A, B, and C, which are entities in the CloudBlock system; the design is deployed at each center to construct a blockchain network. End-users gain trust in this method because the network becomes unchangeable, open to scrutiny, and traceable [80]. When the CloudBlock and EdgeBlock receive the request for information, the preceding protocol is followed, and the service is supplied to the client application, with the administrator being warned in the case of attack cases. The proposed system is based on the combination of blockchain-based on-chain, and AI-based models to create a sustainable smart city.

The proposed framework employed the edge and cloud to take advantage of the combination networks to address the challenges of the IoT-based enabled smart city systems. The effectiveness of the blockchain paradigm was used to protect the IoT-based aided smart cities system, by integrating blockchain and IoT-enabled at edge and cloud sides. Fig. 5.1 displayed the proposed model deployed blockchain and IoT-based enabled edge and cloud called EdgeBlock and CloudBlock. The proposed architecture is to enhanced the smart city applications privacy and security. The suggested framework has three main layers namely: the blockchain



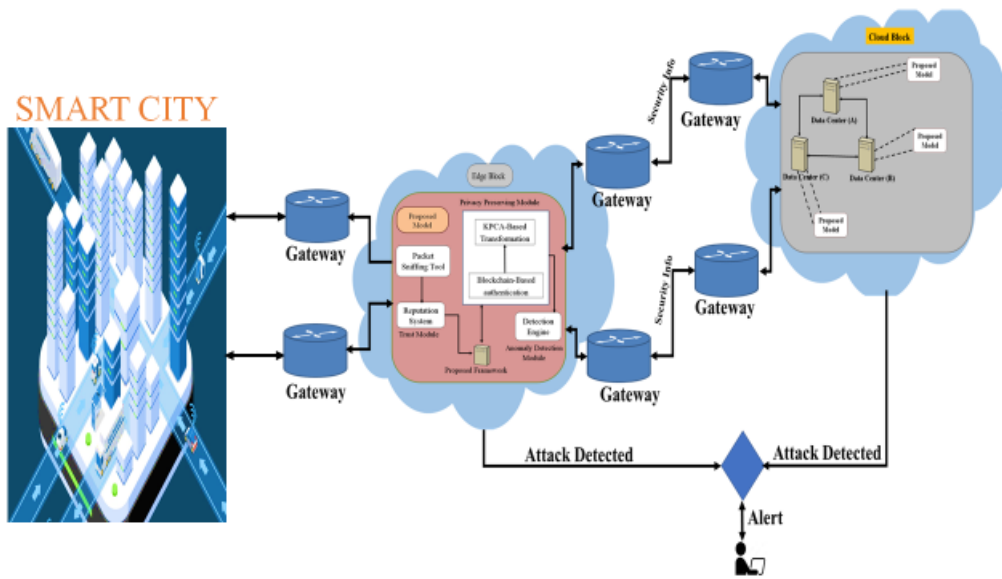


Fig. 5.1: The proposed framework for Privacy and Security Enhancement in Smart Cities.

security management, a two-level privacy-protection technique, and intrusion detection using CNN algorithm as discussed in detail below: Edge placement framework: The nearest gateway/router is used to forward the network traffic to the EdgeBlock once the data is captured and generated from the IoT-based devices within the level of the smart city. To sniff the incoming traffic at the EdgeBlock, the previous involves a packet sniffing device (sensor) to extract the related features. The EdgeBlock is used to bring the processing closer to the client without necessarily needing to work with cloud network nodes.

Address-based blockchain reputation system was used to verify the reliability of the source of captured data and to compute the status score. The transaction is divided into three based on the outcome of the status score: Authentic, Universal, and Deceitful, then stored in the proposed framework. The privacy-protection module stored the three trust data classes alongside the raw data. To generate a message digest with proof of hash in this segment, blockchain-based ePoW was used, and the message summary is distributed into the blockchain system. The next second-level privacy uses the KPCA technique for converting captured data into a newly converted format that prevents intruders and harmful assaults like DoS, FDIA, and DDoS and removes irrelevant parameters through the feature selection. The CNN model is used to classify the data into standard and various types of attack, and the administrator will be alerted based on which the class of the intruders. The security information along with the demand is sent to the CloudBlock module if information available at the edge side is provided for regular contracts where obligatory information is not accessible. CloudBlock Distribution Framework: This consists of various data centers from several vendors. Three centers were used for the proposed model called A, B, and C as displayed in fig.5.1.

The proposed system is installed in each data center, resulting in the formation of a blockchain network. When CloudBlock gets an inquiry, the resource is supplied to the requester, and the administrator is notified of any attack occurrences. Therefore, the proposed system utilized the collaborative combination within edge-cloud infrastructure and integrates Blockchain enabled on-chain, and off-chain with the DL model for the development of a sustainable smart city.

**5.1. Pre-processing.** The preparation stage receives the data set, which consists mostly of two approaches. Data conversion and data normalization are two of them. For processing, the data conversion converts nominal features to numeric features. The data normalization process reduces the huge disparity in attribute values to a reasonable range of values. We employed the minimum-maximum scaling method, which

is formally stated as equation (5.1):

$$Y = \frac{Y - \min(Y)}{(\max(Y) - \min(Y))} \quad (5.1)$$

where  $Y$  denotes the feature value in the data set, and it is in the range of  $[0, 1]$ .

**5.2. Feature Extraction.** Following preprocessing, the dataset is subjected to feature extraction, with KPCA being the most popular method for reducing the data to a lower dimension. As a result, KPCA for complex structures does not take into account non-linear data features. Using KPCA, this issue can be resolved. The feature space  $R$  is represented in the mapping function  $P$  equation (5.2)

$$P : \phi \in R_m \rightarrow P\phi \in R \quad (5.2)$$

where

$$\sum_{i=1}^t P(\phi_i) = 0, \quad (5.3)$$

is the covariance matrix can be calculated using the formula equation (5.4)

$$C_{o_{mtx}} = \frac{1}{t} \sum_{i=1}^t (P(\phi_i) - \text{mean})(P(\phi_i) - \text{mean})^T \quad (5.4)$$

$$M_{ean} = \frac{1}{t} \sum_{i=1}^t P(\phi_i) \quad (5.5)$$

$$C_{o_{mtx}} = \frac{1}{t} \sum_{i=1}^t P(\phi_i)P(\phi_i)^T \quad (5.6)$$

Eigenvalue and Eigenvector are two terms that are used interchangeably. An equation can be used to evaluate equation (5.7)

$$C_{o_{mtx}}I = \lambda_i I \quad (5.7)$$

Combining equations (5.6) and (5.7), we get

$$C_{o_{mtx}} = \frac{1}{t} \sum_{i=1}^t P(\phi_i)IP(\phi_i)^T = \lambda_i I \quad (5.8)$$

The eigenvector can be rewritten using the formula given in equation (5.9)

$$I = \frac{1}{t} \sum_{i=1}^t (\delta_i P(\phi_i)) \quad (5.9)$$

For determining the quotient  $i$ , a kernel matrix  $W$  of size  $t \times t$  is defined. The elements are calculated using equations in this case (5.10).

$$W_{ij} = (P(\phi_i)P(\phi_j)^T = (P(\phi_i) \cdot P(\phi_j) = W(\phi_i, \phi_j) \quad (5.10)$$

when there is no mean in projected dataset  $(P(\phi_i))$ .

**5.3. The Convolutional Neural Network.** The convolution kernel was used to train in the convolution layer for the higher layer's feature map. The outcome is a new feature graph including numerous feature graphs that is fed into a convolution core's input signal. Several feature graphs can be convoluted together to generate another output layer in each output feature graph [46]. The following is how the convolution layer is calculated:

$$X_j^l = f \left( \sum_{i \in M_j} X_i^{l-1} \times K_{ij}^l + b_j^l \right) \quad (5.11)$$

where  $X_j^l$  stands for the  $j$  feature of the layer map  $l$ ,  $K_{ij}^l$  for the convolutional kernel function,  $f$  for the activation function, and  $b_j^l$  and  $M_j$  for the bias parameter and input feature graph respectively. Each output feature graph is formed using a bias coefficient from a combination as an input feature graph. The error signal of the layer with the weights of the feature graphs is determined using the result of the preceding step, and constant  $l$  is set in the bottom sample layer  $\delta$ . The operation is repeated in the convolution layers to obtain the error signal  $b_j^l$  of each feature graph  $j$ .

$$\delta_j^l = \beta_j^{l+1} (f'(u_j^l \cdot up(\delta_j^l + 1)) \quad (5.12)$$

The layer may be used, to sum up the elevation in (5.11), the calculation for a sampling operation is given in equation 5.13:

$$\frac{\delta E}{\delta b_j} = \sum_{u,v} (\delta_j^l)_{uv} \quad (5.13)$$

Finally, the weight gradient of the convolution kernel can be determined using the classic BP approach in CNNs that include weighted values with varied connections. It must first generate a gradient for each link associated with a given weight and then combine the gradients.

$$\frac{\delta E}{\delta K_{ij}^l} = \sum_{u,v} (\delta_j^l)_{uv} (p_i^{l+1})_{uv} \quad (5.14)$$

where  $p_i^{l+1}$  can be multiplied by  $K_{ij}^l$  been a small block element in the convolutional where the value of the output conversion feature graph's is  $(u, v)$  placement. This can be multiplied by a deconvolution element been the result of a small block of the upper  $(u, v)$  position. The lowest sampling layer works based on the notion that each outcome featured chart is a small representation of the convolution layer.

$$X_j^i = f(\beta_j^i \text{down}(X_j^{l-1}) + b_j^i) \quad (5.15)$$

The  $n$  times smaller for feature graph to achieve scaling invariance, where  $\text{down}(X_j^{l-1})$  is the specimen frame and the low bit value is  $n * n$ . Each output feature graph has its own multiplying offset variable and admixture bias variable  $\beta$ .

$$\delta_j^l = f'(u_j^i) \cdot \text{conv2}(\delta_j^{l+1}, \text{rot180}(k_j^l - 1), \text{"full"}) \quad (5.16)$$

To connect the convolution function to the entire convolution function, and before calculating, the volume kernel must be rotated 180 degrees. In complement 0, it can handle the convolution border as well as the missing pixel. After that, a  $t_0$  will be obtained as equation 5.17:

$$\frac{\delta E}{\delta b_j} = \sum_{u,v} (\delta_j^l)_{uv} \quad (5.17)$$

$$\frac{\delta E}{\delta \beta_j} = \sum_{u,v} (\delta_j^l \cdot \text{down}(X_j^{l-1}))_{uv} \quad (5.18)$$

Table 5.1: Characteristics of the ToN-IoT Dataset

Class	Total	Characteristics
Benign	300,000	Standard unmalicious movements
Backdoor	20,000	A way of exploiting remote devices by reacting to client applications that have been carefully designed.
DoS	20,000	An effort to overwhelm a workstation system's resources in order to obstruct access to its data.
DDoS	20,000	A similar approach to DoS, but with several scattered sources.
Injection	20,000	SQL injection and code injection are some of the most prevalent assaults that use unverified inputs to modify the course of operation.
MITM	1043	Person in the Middle is a technique of eavesdropping on traffic and conversations that involves putting an assailant between a target and the host with which the target is seeking to contact.
Password	20,000	includes a variety of brute-force and sniffer methods aimed at collecting credentials.
Ransomware	20,000	An assault in which data on a server are encrypted and money is demanded in return for the decoding technique.
Scanning	20,000	An exploit that encrypts data on a website and requests money in exchange for the decoding process.
XSS	20,000	Cross-site Scripting (XSS) is a type of infiltration in which the intruder transmits malicious files to end users via internet apps.

In the convolution neural net, the frequency increase from time  $t$  to time  $t + 1$  is analogous to the BP approach.

$$w(t + 1) = w(t) + \mu\delta(t)x(t) \quad (5.19)$$

Here  $\delta(t)$  is the error term, the learning rate is  $\mu$ , and the input of the neuron is denoted as  $x(t)$ .

**5.4. Datasets.** Two prominent datasets that were publicly available were used in the study namely: ToN-IoT and BoT-IoT datasets.

**5.4.1. The ToN-IoT Dataset.** The dataset is freely available established by UNSW Canberra Cyber IoT-Lab at The Australian Defense Force Academy, and is gotten from a practical and large-scale system [26-27]. A variety of normal and cybersecurity Incidents from IoT networks are compiled in parallelization for the dataset.

To replicate the capability as well as the adaptability of automotive IoT and industry Networks 4.0, the IoT lab has constructed a new testbed that connects Simulated machines, physical equipment, hacking platforms, and cloud and fog systems are all examples of technology, and IoT sensors are among the devices available. The dataset covers many modern DoS, DDoS, and other device-connected assaults spyware, that have been installed via the IoT network in comparison to web apps, Internet of Things interfaces, and electronic systems. There are 43 features in the dataset. The dataset was split into two parts: a train set and a test set, each comprising 70% and 30% of the total. The statistics of the collection containing both standard and unique network attacks are shown in table 5.1.

**5.4.2. The BoT-IoT Dataset.** This data was developed by creating a realistic network platform in the UNWS of Canberra Cyber Range Center [32]. The Message Queuing Telemetry Transport protocol is used to generate this data gathering, which connects machine-to-machine connections, making it a viable alternative for IoT solutions. Table 1 shows the statistics of various assaults in the BoT-IoT dataset. Various assaults, like DDoS, DoS, and Theft, are included in the dataset. Table 5.2 shows the characteristics of the BoT-IoT dataset.

Table 5.3 shows the reduced features after applying the KPCA algorithms to the datasets used to assess the effectiveness of the proposed approach. The features of the ToN-IoT were reduced to 17 features while BoT-IoT features were reduced to 7 features.

**5.5. Performance analysis.** The accompanying performance indicators were used to test the hypothesized algorithm's results and evaluate it to other latest systems based on DL and hybrid rule-based techniques. The amount of right/wrong outputs in a classifying job was totaled and evaluated to the benchmark findings. Accuracy, Precision, Sensitivity, Specificity, and F1-score are the most commonly used matrices. The numerical

Table 5.2: Features of the used BoT-IoT Dataset

Class	Total	Characteristics
Benign	477	Normal unmalicious flows
Reconnaissance	91,082	A way of exploiting remote devices by reacting to client applications that have been carefully designed.
DoS	1,650,260	An effort to overwhelm a computer device's resources in order to obstruct accessibility to its information.
DDoS	1,926,624	The DDoS is a type of DoS attack that uses many distributed sources.
Theft	79	Data theft and keylogging are examples of attacks aimed at obtaining sensitive data.

Table 5.3: The selected features for ToN-IoT and BoT-IoT Datasets using the KPCA feature selection method.

Class	Total Features Selected	Selected Features
Benign	17	F1, F2, F3, F4, F5, F6, F7, F8, F9, F10, F11, F12, F13, F14, F15 F16, F17
Recon	7	F1, F2, F3, F4, F5, F6, F7

metrics true positive (TP), true negative (TN), false positive (FP), and false-negative (FN) were obtained to rectify the confusion matrix, as shown in equations 5.20 - 5.26 [81].

$$Accuracy = \frac{(TP + TN)}{(TP + FP + FN + TN)} \quad (5.20)$$

$$Precision = \frac{TP}{(TP + FP)} \quad (5.21)$$

$$Sensitivity \text{ or } Recall = \frac{TP}{(TP + FN)} \quad (5.22)$$

$$Specificity = \frac{TN}{(TN + FP)} \quad (5.23)$$

$$F1 - score = \frac{2 * (Precision * Recall)}{(Precision + Recall)} \quad (5.24)$$

$$TPR = \frac{TP}{(TP + FN)} \quad (5.25)$$

$$FPR = \frac{FP}{(FP + TN)} \quad (5.26)$$

**6. Results and Discussion.** The platforms used for the experiment and implementation of the proposed models are the R programming language and the Scikit-learn library used for the implementation of the AI model [82]. The KPCA was used for feature selection in order to eliminate the unnecessary elements from both datasets, and the suggested model was evaluated using performance indicators. The Solidity programming language with Ethereum was used for the implementation of the Blockchain Machinery. An HP with Window 11 Intel(R) Core(TM) i7-2520M CPU @ 2.50GHz with 128 GB RAM and 2 TB hard disk. The results were compared with the recent state-of-the-art model using the same datasets, non-blockchain and blockchain framework.

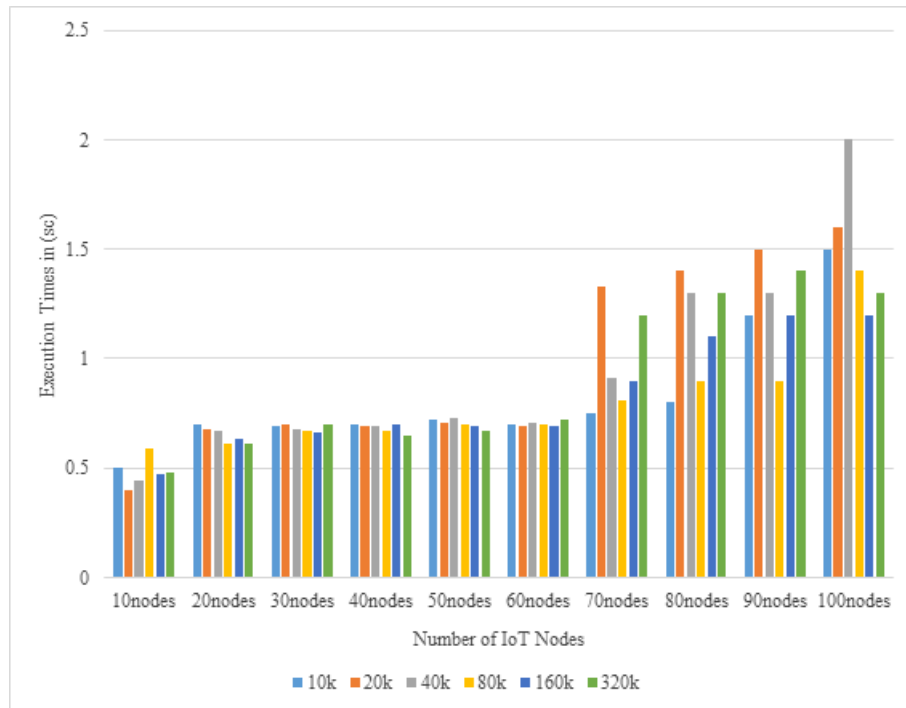


Fig. 6.1: The effect of adjusting the amount of IoT nodes in the CloudBlock off-chain memory in a Blockchain network on posting duration.

**6.1. The proposed model for security and privacy process.** The IoT-based devices' dependability and reliability were evaluated, and the CloudBlock was used to the data for general transactions and honesty. The proposed model was studied in uploading files on the CloudBlock storage. The peers are represented from 10-100 in the X-axis, and the time taken to upload various file sizes is represented in the Y-axis with sizes of 10KB-30Kb. In the off-chain storage network. Fig.6.1 shows the results obtained using the method and Investigates the time impact when the system is expanded statically. There is a direct link between the quantity of active participants and the memory usage that grows as more IoT units are joined to the network must be published in terms of scalability.

**6.2. Trust Management Process.** The reputation score was computed for the proposed model using 100 nodes of IoT in the Ethereum network, and the IoT nodes were represented by  $R_{ps}$ . The specific address was allotted to each IoT node in the blockchain network. The transaction score was computed using the transactions performed by these nodes, using  $T_{xscore}$  against the  $C_{Threshold}$  confidence threshold value. To evaluate the trust within the proposed model, the score generated was used by calculating  $R_{ps}$  for produced transactions. The  $R_{ps}$  and  $T_{xscore}$  for 1000 transactions are displayed in Fig. 6.1.

The valid transaction values for the available features were used in the model to calculate for ToN-IoT and BoT-IoT datasets, hence, fall in the honest operation grouping. The proposed technique was used to study the execution time-varying number of transactions with respect to computation in the CloudBlock framework. The results show that there are steady increases in the  $T_{xscore}$  as the number of transactions increases. Fig.6.2 shows the IoT device trust management using reputation score vs. transaction score.

Fig.6.3 displays the results of varying IoT nodes with various sizes in block creation time using the proposed model. Chronologically ordered the blocks are added to the distributed ledger. Once the collaborating IoT nodes have completed the mining operation. The message digest is published throughout the blockchain network, along with the reputation score. The verifiability and immutability of the transactions signify these chains of blocks. The accessibility of the operation in the blockchain system was shown using block access time. The

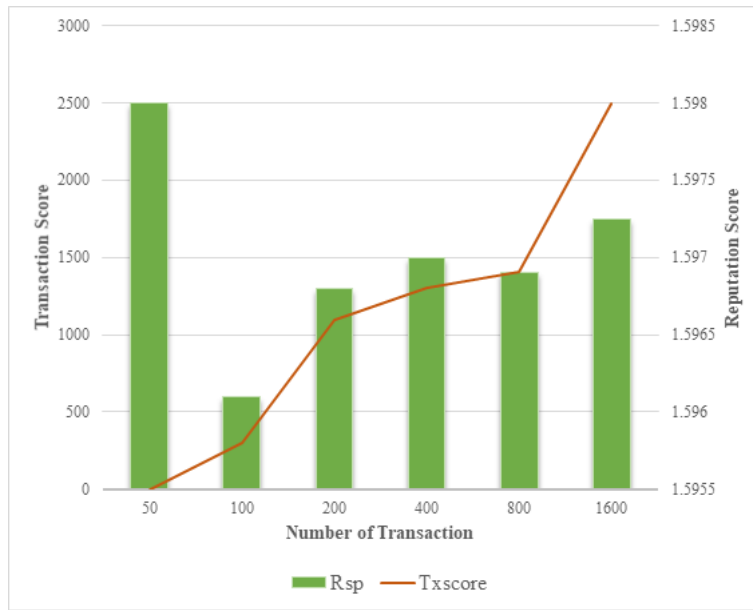


Fig. 6.2: The IoT device Trust Management using Reputation score vs. Transaction score

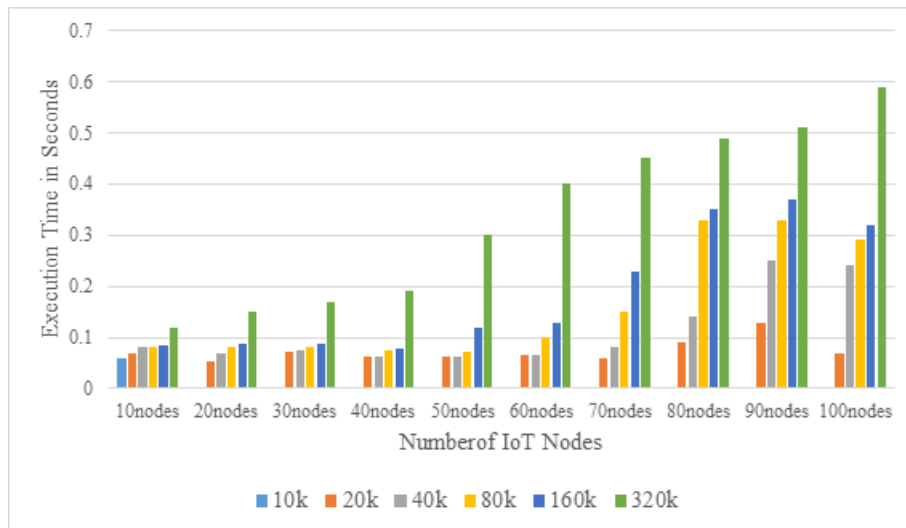


Fig. 6.3: The effects of file size and the number of IoT nodes in a blockchain network on block

block access time of the transaction in the blockchain is shown in fig.6.4. The access time of the block for the proposed framework was calculated by varying the number of IoT nodes for KPCA evaluation. The results show that an increase in the number of IoT nodes also rises the access time taken in the network.

**6.3. The intrusion detection process experiments results.** The performance of the suggested two-level security and confidentiality method is assessed using two prominent IoT smart cities datasets as a function scheme of the intrusion detection based CNN and KPCA was used for features selection. The proposed model was applied to the two datasets in other to be able to identify various attack instances and normal.

The proposed model uses various performance metrics for evaluation. Preprocessing was performed utilizing

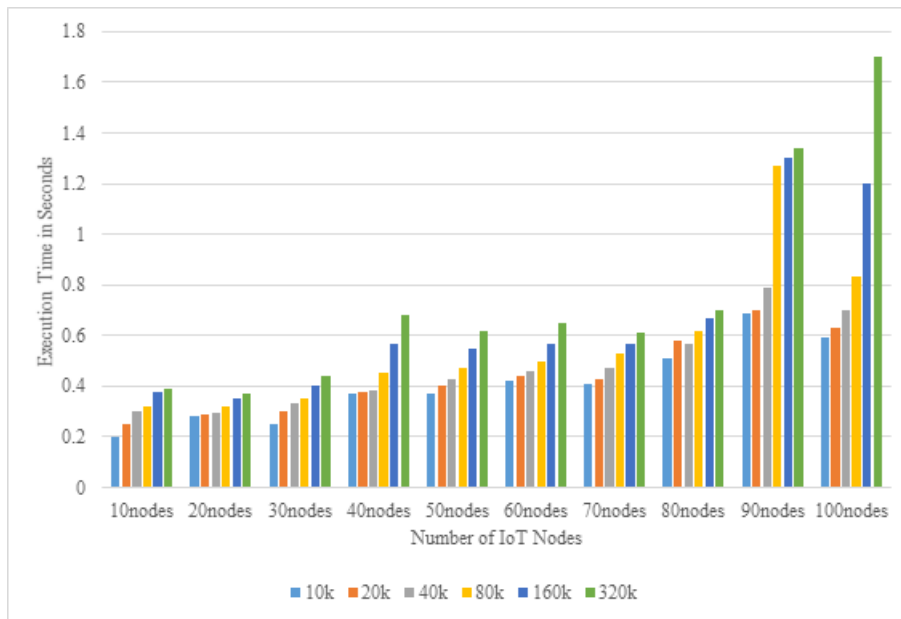


Fig. 6.4: Block access time in a blockchain network with different file sizes and numbers of IoT nodes.

feature mapping, variable selection employing KPCA, and normalized on both ToN-IoT and BoT-IoT datasets. The most relevant features from the used datasets were selected using the KPCA algorithm. Table 6.1 shows the list of the most relevant features selected for the ToN-IoT dataset in the designed security and privacy models using the CNN algorithm for the smart city.

Table 6.1 shows the detection rate (DR) for the instances in the ToN-IoT using the proposed model. The detection rate of the recorded instances with feature selection is as follows: Benign (100.0%), Backdoor (100.0%), Dos (98.7%), DDoS (97.4%), Injection (98.3%), MiTM (69.2%), Password (98.2%), Ransomware (99.5%), Scanning (99.3%), and XSS (97.2%), respectively. The DR of the recorded instances without feature selection is as follows: Benign (99.30%), Backdoor (99.9%), Dos (98.7%), DDoS (91.9%), Injection (92.4%), MiTM (60.6%), Password (95.8%), Ransomware (98.3%), Scanning (93.4%), and XSS (80.1%), respectively. The DR results for the BoT-IoT dataset using the proposed model are shown in Table 6.2 to determine the performance of the models on the detection of attackers on the dataset without feature selection on class types for Benign (78.0.0%), DoS (99.7%), DDos (100.0%), Reconnaissance (100.0%), and Theft (100.0%), respectively. The DR of the recorded instances without feature selection are as follows: Benign (82.6%), DoS (90.2%), DDos (95.7%), Reconnaissance (97.4%), and Theft (0.0%), respectively. The results proposed technique demonstrated high DR for intrusions detection in both datasets and show better performance when compared with recent models. The obtained resulted from the two datasets show that the feature selection technique really works effectively on the datasets for the detection of intrusions attack.

Table 6.3 shows the proposed model performance using both ToN-IoT and BoN-IoT datasets with the aforementioned metrics. The results using the model are very relevant and effective in intrusion detection of IoT-based enabled smart cities for the classification and prediction of various attacks with the nodes of the networks.

From the results obtained It can be concluded that the proposed system's performance is satisfactory for the security and privacy of the IoT-based enabled smart city looks promising and work perfectly, especially when the KPCA feature selection is applied to the datasets. The use of blockchain for first-level protection before applying the DL model with KPCA performs very well better than the existing model without the application of blockchain and feature selection algorithms.



Table 6.1: Detection rates for ToN-IoT datasets with and without Feature Selection

Attack Class	Dataset Used	
	With Feature Selection	Without Feature Selection
Benign	100.0	99.3
Backdoor	100.0	99.9
DoS	98.7	98.7
DDoS	97.4	91.9
Injection	98.3	92.4
MITM	69.2	60.6
Password	98.2	95.8
Ransomware	99.5	98.3
Scanning	99.3	92.4
XSS	97.2	80.1

Table 6.2: Detection rates for BoT-IoT datasets with and without Feature Selection

Attack Class	Dataset Used	
	With Feature Selection	Without Feature Selection
Benign	78.0	82.6
DoS	99.7	90.2
DDoS	100.0	95.7
Reconnaissance	100.0	97.4
Theft	100.0	0.0

Table 6.3: The Proposed Model Evaluation Results.

Dataset	TP Rate	FP Rate	Precision	Sensitivity	F-Measure	ROC	Class
ToN-IoT	99.8	0.1	97.6	99.9	98.9	98.9	Attack
	100.0	0.001	99.8	99.6	97.5	99.8	Normal
BoN-IoT	100.0	0.1	99.3	99.9	99.7	99.7	Attack
	99.9	0.001	100.0	99.3	99.2	99.8	Normal

**6.4. Using Existing Techniques to Compare the Proposed Model.** Table 6.4 shows the contrast results of the anticipated models with some current work to really show the efficiency of the KPCA feature selection on the datasets and the classifier detection efficiency. To demonstrate how feature selection influences the discovery performance of a classification procedure, the outcomes show that the suggested model performs better in precision and FPR than other approaches.

The proposed model in the overall accuracy on both datasets has a 99% and the FPR is a very low error percentage. The suggested model also shows a better performance across all evaluation metrics used. The marginal increases in the accuracy may be due to the use of the KPCA algorithm in selecting the most appropriate features for each dataset.

To further prove the efficiency and performance of the proposed framework, a comparison of both non-blockchain and blockchain structures was conducted with other related schemes. These comparisons were based on various parameters like (i) off-chain, (ii) ledger distribution, (iii) IDS, (iv) security, (v) privacy, (vi) trust, (vii) scalability, (viii) decentralized (ix) non-repudiation, (x) verifiability, and (xi) deep learning. The scalability of the system is achieved with the off-chain storage, their blockchain memory allocated hashing used for each raw transaction record. The ledger distribution is the capacity of each IoT node to save the matching replica of the ledger and replicate itself in the IoT-based enabled smart city. framework. The IDS was achieved with the use of the AI-based model classifier.

The two-level approach for security and privacy preservation used by the framework addressed the major challenges of security, privacy, and trust in the smart city. The combination of the blockchain paradigm, KPCA algorithm, and AI classifier security and privacy model ensures these features.

Table 6.4: The UNSW-NB15 Dataset Performance Comparison Summary

Model	Performance Metrics					
	Accuracy	FPR	F-Score	Sensitivity	Precision	ROC Curve
Wrapper + neurotree [83]	98.38	1.62	98.4	98.0	98.9	99.8
SVM+EML+K-Means [84]	95.75	1.87	94.4	99.7	89.7	98.6
GA +SVM [85]	97.3	0.017	96.6	99.7	93.8	98.1
CNN+LSTM [86]	94.12	-	95.6	98.9	92.5	98.4
Modified KNN [74]	98.7	1.3	99.2	99.6	98.8	99.8
CfsSubsetEval + GA+RuleEval+ANN [38]	98.8	1.2	98.9	98.9	98.9	99.8
DFFN + Rule-based [20]	98.9	1.1	98.9	99.8	96.7	98.9
Proposed Model	99.8	1.07	99.2	99.9	100	99.8

Table 6.5: Comparison of the proposed model with existing blockchain models

Model	1	2	3	4	5	6	7	8	9	10	11
RealAlert [28]	x	x	x	√	x	√	x	x	x	x	x
Cryptographic + blockchain [87]	x	√	x	x	√	X	x	√	√	√	x
Fog + Blockchain [88]	x	√		√	√	√		√		√	X
Interplanetary file system + Blockchain [89]	√	√	x	√	√	X	√	√	√	√	x
TP2SF [1]	√	√	√	√	√	√	√	√	√	√	x
TrustChain [67]	x	√	x	√	√	√	x	√	x	√	x
BiLSTM + Blockchain [71]		√	√	√	√			√		√	√
PPAD-CPS [90]	x	x	√	√	√	X	x	√	x	x	√
Proposed Model	√	√	√	√	√	√	√	√	√	√	√

Table 6.6: Comparison of the proposed model with non-blockchain models

Model	1	2	3	4	5	6	7	8	9	10	11
independent component analysis (ICA) [91]	x	x	√	√	√	X	x	x	x	x	x
SUM, RNN and LSTM [32]	x	x	√	√	x	X	x	x	x	x	√
CST-GR [92]	x	x	√	√	x	X	x	x	x	x	x
TP2SF [93]	√	√	√	√	√	√	√	√	√	√	x
DFFN + Rule-based [20]	x	x	√	√	x	X	x	x	x	x	√
PCA-firefly + XGBoost [94]	x	x	√	√	x	X	x	x	x	x	x
FGMC-HADS [33]	x	x	√	√	x	X	x	x	x	x	√
bijjective soft set [95]	x	x	√	√	x	X	x	x	x	x	x
bijjective soft set + CorrACC [96]	x	x	√	√	x	X	x	x	x	x	x
Proposed Model	√	√	√	√	√	√	√	√	√	√	√

The proposed framework ensures the maintainability of non-repudiation with recorded transactions in the ledger since the participant node cannot be denied from a transaction once the transaction gets included in the ledger, and the behavior of the transaction is been protected. The use of timestamp records to audit-trail each transaction proved the verifiability metric in the framework. The CloudBlock framework with EdgeBlock ensured decentralization in the IoT-based enabled smart city. The results of both blockchain and non-blockchain for comparison are presented in Tables 6.5 and 6.6.

A distributed network built on the blockchain is being created to secure transactions in IoT-based smart cities by confirming the transaction's validity. Blockchain technology is used to transfer data between locations. The blockchain functions in a number of ways, including block production, block broadcast to all network nodes, transaction verification, and transfer request. The transaction will continue with the addition of a new user registration if the verification process is successful; else, all steps must be redone. The fundamental purpose of trusted entities is to validate transactions; hence this is how they are employed. After the transaction has been validated, the verifier must choose between the options yes or no. If the state is accurate, the transaction is finished by adding a block to the Blockchain; if an attack is discovered during the transaction, it is aborted.

The sort of assaults can be identified and the transaction can be blocked by utilizing a deep learning-based classifier that is enabled with KPCA. The dataset's features are first extracted using KPCA, and then the extracted features are submitted to CNN for classification.

The creation of an improved Proof of Work (ePoW) method based on blockchain is part of the first level of privacy and trust, and it will be used to authenticate data records and stop data poisoning assaults from changing the original data. The attribute data is subjected to the second level of privacy while the ePow algorithm is running. The datasets for smart cities are referred to by this property. At this level, we choose crucial features in order to safely train and validate a utility model. Utilizing a set of weighted parameters and a feature selection model called a KPCA, an input  $X$  is encoded into new data codes.

The proposed framework was able to secure and preserve IoT-based SC with given trust among users and identified attack behaviors efficiently. This can be ascribed to its layered data extraction of the IDS model development. The two-level security and privacy technique of blockchain with smart contract help in achieving better protection against intruders through the validation of data transactions, validation of the IDS model, and the extraction of relevant features from the captured data for processing purposes. The first phase makes use of security and privacy-preserving based on blockchain technology to insured data integrity by verifying and checking records for possible poisoning through the application of hash chain making malevolent modification of records highly exclusive computationally. The second phase makes use of KPCA for feature selection, which includes data preprocessing, converting nominal attributes into numeric using feature mapping, and scaling the attributes into specific ranges by feature normalization. The two-level security and privacy-preserving technique protect sensitive data and information of smart cities nodes network traffic and power systems against exposure in IoT-based applications. The incorporation of CNN and KPCA in the second stage fuses important features into relevant features to be able to discover intrusion attacks in a smart city classifying malicious activities with the IoT nodes in SC using lower and upper boundaries of the normal subsequent likelihoods, and the variations from them should be treated as abnormal discoveries.

The proposed model is very easy to deploy, implement, and in an extremely active and heterogamous network of IoT-based enabled smart cities, the system can be efficiently used to detect recent and most attacks found within network nodes and protect and secure the IoT-based enabled smart city. The proposed framework can competently compute the status score of the participating IoT nodes using the blockchain reputation system and build trustworthiness in the network traffic and transmission. The integration of blockchain, KPCA and CNN approaches for security-privacy preserving, and the dimensionality reduction method helps in preventing inference and poisoning attacks within IoT nodes networks. Therefore, the proposed model improved the overall efficiency and performance of the IoT-based enabled smart city in general. The verifiability, reliability, and traceability were able to be activated with the incorporation of CloudBlock a blockchain paradigm cloud infrastructure. The use of DL algorithms with feature selection techniques also increases the performance of the anticipated structure greatly.

**6.5. The benefits and challenges related to deploying the CloudBlock-EdgeBlock architecture in a real-world smart city environment.** The proposed architecture has a lot of benefits. The majority of assaults encountered in the highly dynamic and heterogeneous network of an IoT-driven smart city can be efficiently detected by it, and it is first straightforward to implement and deploy. Second, the network becomes more trustworthy thanks to the address-based blockchain reputation system's ability to accurately calculate the reputation score of participating IoT nodes. Third, the two-level privacy-preserving strategy combines PCA-based dimensionality reduction with blockchain technology to thwart inference and poisoning attacks. The performance of the suggested TP2SF architecture has been significantly enhanced by the aforementioned method. Additionally, the incorporation of blockchain technology into cloud and fog infrastructures allows for verifiability, traceability, and dependability. However, a few issues have been noted, such as the fact that block mining and file uploading times gradually increase as more IoT nodes participate.

A transaction in the blockchain is only confirmed after being approved by/verified by all nodes. There is a risk for cyberattacks because this verification takes a certain amount of time. One of these attacks that takes advantage of the transaction verification time is double-spending. Attackers take advantage of the time required for the authentication of each transaction on the blockchain. The attacker uses the same coin twice during the transaction verification delay since both transactions are being verified at the same time. This

makes it simple to copy and fake digital currencies. The blockchain's immutability guarantees data integrity, facilitates message exchanges between all parties, and creates logs and events. It ensures that everyone has access to the deployed smart contract on the blockchain at all times. Additionally, availability guarantees that all services are constantly accessible. The system is also shielded against DoS assaults because to the fact that all transactions are recorded on an Ethereum distributed ledger. So there is no concern about hacking, failure, or compromise. Thousands of reliable mining nodes guard the Ethereum ledger, making it extremely resistant to DoS attacks. Using a permissioned or private blockchain, such as Hyperledger or private Ethereum networks, the criterion of confidentiality is met. In the suggested situation, the proposed solution is built on a blockchain with a permissions network.

Links between transactions and public blockchain addresses can reveal a user's true identity. Systems for managing digital identities centrally are not secure. A new contract deployment for every upgrade raises trust and inconsistency issues. The smart contract is unable to send deterministic outside requests. Smart contracts are unchangeable and aid in building confidence between the parties to a contract. The smart contract code, however, is typically not upgradeable, even in the event of flaws, vulnerabilities, or new business-logic requirements (for example, on the Ethereum Platform). A new instance of the smart contract with a new contract address is typically used to deploy an updated smart contract code, which may cause issues with inconsistency. Delegating calls from the proxy contract to the new logic contract is one workaround for an upgradeable smart contract, nevertheless [5]. While the logic contract performs the new logic, the proxy contract holds the data. With each change, the logic-contract address is updated in the proxy contract. The upgrade has no impact on the smart city service user because the proxy contract protects his data. However, there are trust and decentralization problems with the proxy-contract-delegation-call approach.

Since everyone can typically access user data and user-pseudonymous identities on public blockchains, privacy issues, and identity risks arise. Innovative privacy techniques including double-blind data sharing and zero-knowledge proofs-based distributed permission management can be used for privacy-preserving selective data sharing in mutually anonymous multi-party transactions in a variety of smart city services, according to authors in [19]. While the transaction data itself may be encrypted via symmetric on-chain encryption and other methods. One of the drawbacks is that each of these methods increases network latency.

Nowadays, digital identity management systems run by centralized authorities are used to give user identities for smart city services. Users can completely control their digital identity without the help of a centralized third party using self-sovereign identity (SSI) and decentralized ID (DID). Users can manage how their individually identifiable information and data are shared in this way. Blockchain-enabled SSI and DID can be utilized for decentralized user identification, authentication, and authorization in IoT-enabled smart city services. However, there are a number of problems with SSI and DID, including human reliance (a user could misplace the private key). A significant difficulty is creating safe recovery strategies for SSI and DID.

Every transaction on open blockchains is viewable by anybody. The public address of each participating device can be used to identify it. Even though the public address is fictitious, curious, or malicious individuals with knowledge of the background can take advantage of the connections between public addresses and the real-world identities of the transaction users. The privacy concerns in applications for smart cities powered by cryptocurrencies can be reduced by creating a new disposable address for each new payment as well as by deploying mixers that gather and disperse funds to pertinent stakeholders.

**7. Conclusion.** The increased use of the internet in recent years globally has made it a favorable environment for nefarious activity. Malicious IoT-based systems for SCs are an example of these activities, and they are one of the greatest severe pressures to internet users and smart cities. This rapid development has dramatically increased the number of urbanities that moved to the internet to create a smart city. The fast-growing in these areas expanded online crimes that run using malicious network nodes. Detection and analysis of these networks became one of the major problems of online space and this is evident in the evolution of SCs. To alleviate the challenges associated with IoT-based enabled SCs like security, privacy, scalability, centralization, and communication latency. This paper proposes a blockchain, KPCA with DL-based technique-oriented infrastructure for a secure IoT-based enabled SC. The communication stage of the IoT-based smart city was secure and established protocols for data forwarding using blockchain technology in the IoT nodes network. The KPCA was used to eradicate unwanted and unrelated features from the dataset. The communication

latency, scalability, and the detection of an intrusion were achieved using a DL-based cloud employed at the application level. The suggested model was assessed utilizing two openly accessible transformed and original datasets namely: ToN-IoT and BoN-IoT for the security mechanism since both datasets have various IoT-based attacks like DoS, Ransomware, normal vectors, MITM, and Theft as mentioned earlier, and are publicly available. The results show a higher level of performance in terms of accuracy and increase the security, privacy, and maintainability concerns of IoT-based enabled SCs applications. In terms of accuracy, recall, precision, detection rate, and F1-score, the proposed systems' experimental results outperform other existing state-of-the-art models when compared. Both blockchain and non-blockchain systems demonstrated the advantage of the proposed system against existing frameworks. Finally, the proposed system enabled IoT-based enabled smart city high-performance computing resources, and cost-effective applications like smart waste management, smart grid, smart agriculture, and smart healthcare. The future scope will extend the proposed system by integrating cryptography techniques that will further secure the cloud from cyberattacks and give privacy to the users' data. The implementation of a protocol that will verify the overall system to recognize security restrictions. These will further enhance the overall security and privacy necessity in SC. The utilization of the anticipated model increases automatic data analysis and the advanced communication bandwidth of smart cities.

## REFERENCES

- [1] Park, E., Del Pobil, A. & Kwon, S. The role of Internet of Things (IoT) in smart cities: Technology roadmap-oriented approaches. *Sustainability*. **10**, 1388 (2018)
- [2] Zhang, X. The trends, promises and challenges of urbanisation in the world. *Habitat International*. **54** pp. 241-252 (2016)
- [3] Singh, S., Sharma, P., Yoon, B., Shojafar, M., Cho, G. & Ra, I. Convergence of blockchain and artificial intelligence in IoT network for the sustainable smart city. *Sustainable Cities And Society*. **63** pp. 102364 (2020)
- [4] Tabane, E., Ngwira, S. & Zuva, T. Survey of smart city initiatives towards urbanization. *2016 International Conference On Advances In Computing And Communication Engineering (ICACCE)*. pp. 437-440 (2016)
- [5] Zhuang, P., Zamir, T. & Liang, H. Blockchain for cybersecurity in smart grid: A comprehensive survey. *IEEE Transactions On Industrial Informatics*. **17**, 3-19 (2020)
- [6] Arowolo, M., Ayegba, P., Yusuff, S. & Misra, S. A Prediction Model for Bitcoin Cryptocurrency Prices. *Blockchain Applications In The Smart Era*. pp. 127-146 (2022)
- [7] Awotunde, J., Folorunso, S., Ajagbe, S., Garg, J. & Ajamu, G. AiIoMT: IoMT-based system-enabled artificial intelligence for enhanced smart healthcare systems. *Machine Learning For Critical Internet Of Medical Things: Applications And Use Cases*. pp. 229-254 (2022)
- [8] Othman, S., Almalki, F., Chakraborty, C. & Sakli, H. Privacy-preserving aware data aggregation for IoT-based healthcare with green computing technologies. *Computers And Electrical Engineering*. **101** pp. 108025 (2022)
- [9] Khatua, P., Ramchandaramurthy, V., Kasinathan, P., Yong, J., Pasupuleti, J. & Rajagopalan, A. Application and assessment of internet of things toward the sustainability of energy systems: Challenges and issues. *Sustainable Cities And Society*. **53** pp. 101957 (2020)
- [10] Silva, B., Khan, M. & Han, K. Towards sustainable smart cities: A review of trends, architectures, components, and open challenges in smart cities. *Sustainable Cities And Society*. **38** pp. 697-713 (2018)
- [11] Nayyar, A., Rameshwar, R. & Solanki, A. Internet of Things (IoT) and the digital business environment: a standpoint inclusive cyber space, cyber crimes, and cybersecurity. *The Evolution Of Business In The Cyber Age*. **10** pp. 9780429276484-6 (2020)
- [12] Ogundokun, R., Arowolo, M., Misra, S. & Awotunde, J. Machine learning, IoT, and blockchain integration for improving process management application security. *Blockchain Applications In The Smart Era*. pp. 237-252 (2022)
- [13] Ogundokun, R., Arowolo, M., Misra, S. & Damasevicius, R. An Efficient Blockchain-Based IoT System Using Improved KNN Machine Learning Classifier. *Blockchain Based Internet Of Things*. pp. 171-180 (2022)
- [14] Chaabouni, N., Mosbah, M., Zemmari, A., Sauvignac, C. & Faruki, P. Network intrusion detection for IoT security based on learning techniques. *IEEE Communications Surveys and Tutorials*. **21**, 2671-2701 (2019)
- [15] Khan, K., Mehmood, A., Khan, S., Khan, M., Iqbal, Z. & Mashwani, W. A survey on intrusion detection and prevention in wireless ad-hoc networks. *Journal Of Systems Architecture*. **105** pp. 101701 (2020)
- [16] Awotunde, J., Chakraborty, C. & Folorunso, S. A secured smart healthcare monitoring systems using Blockchain Technology. *Intelligent Internet Of Things For Healthcare And Industry*. pp. 127-143 (2022)
- [17] Awotunde, J., Jimoh, R., Folorunso, S., Adeniyi, E., Abiodun, K. & Banjo, O. Privacy and security concerns in IoT-based healthcare systems. *The Fusion Of Internet Of Things, Artificial Intelligence, And Cloud Computing In Health Care*. pp. 105-134 (2021)
- [18] Da Costa, K., Papa, J., Lisboa, C., Munoz, R. & Albuquerque, V. Internet of Things: A survey on machine learning-based intrusion detection approaches. *Computer Networks*. **151** pp. 147-157 (2019)
- [19] Bhaskaran, K., Ilfrich, P., Liffman, D., Vecchiola, C., Jayachandran, P., Kumar, A., Lim, F., Nandakumar, K., Qin, Z., Ramakrishna, V. & Others Double-blind consent-driven data sharing on blockchain. *2018 IEEE International Conference On Cloud Engineering (IC2E)*. pp. 385-391 (2018)

- [20] Awotunde, J., Chakraborty, C. & Adeniyi, A. Intrusion detection in industrial internet of things network-based on deep learning model with rule-based feature selection. *Wireless Communications And Mobile Computing*. **2021** pp. 1-17 (2021)
- [21] Hassan, M., Rehmani, M. and Chen, J. Privacy preservation in blockchain based IoT systems: Integration issues, prospects, challenges, and future research directions. *Future Generation Computer Systems*. **97** pp. 512-529 (2019)
- [22] Elrawy, M., Awad, A. & Hamed, H. Intrusion detection systems for IoT-based smart environments: a survey. *Journal Of Cloud Computing*. **7**, 1-20 (2018)
- [23] Singh, S., Sharma, P., Moon, S., Moon, D. & Park, J. A comprehensive study on APT attacks and countermeasures for future networks and communications: challenges and solutions. *The Journal Of Supercomputing*. **75** pp. 4543-4574 (2019)
- [24] Bostami, B., Ahmed, M. & Choudhury, S. False data injection attacks in internet of things. *Performability In Internet Of Things*. pp. 47-58 (2019)
- [25] Illiano, V. & Lupu, E. Detecting malicious data injections in wireless sensor networks: A survey. *ACM Computing Surveys (CSUR)*. **48**, 1-33 (2015)
- [26] Altaf, A., Abbas, H., Iqbal, F. & Derhab, A. Trust models of internet of smart things: A survey, open issues, and future directions. *Journal Of Network And Computer Applications*. **137** pp. 93-111 (2019)
- [27] Awotunde, J., Bhoi, A. and Barsocchi, P. Hybrid cloud/Fog environment for healthcare: an exploratory study, opportunities, challenges, and future prospects. *Hybrid Artificial Intelligence And IoT In Healthcare*. pp. 1-20 (2021)
- [28] Mohiyuddin, A., Javed, A., Chakraborty, C., Rizwan, M., Shabbir, M. & Nebhen, J. Secure cloud storage for medical IoT data using adaptive neuro-fuzzy inference system. *International Journal Of Fuzzy Systems*. **24**, 1203-1215 (2022)
- [29] Behera, S., Pradhan, A. & Dash, R. Deep Neural Network Architecture for Anomaly Based Intrusion Detection System. *2018 5th International Conference On Signal Processing And Integrated Networks (SPIN)*. pp. 270-274 (2018)
- [30] Dash, S., Chakraborty, C., Giri, S., Pani, S. & Frnda, J. BIFM: Big-data driven intelligent forecasting model for COVID-19. *IEEE Access*. **9** pp. 97505-97517 (2021)
- [31] Farnaaz, N. & Jabbar, M. Random forest modeling for network intrusion detection system. *Procedia Computer Science*. **89** pp. 213-217 (2016)
- [32] Koroniotis, N., Moustafa, N., Sitnikova, E. & Turnbull, B. Towards the development of realistic botnet dataset in the internet of things for network forensic analytics: Bot-iot dataset. *Future Generation Computer Systems*. **100** pp. 779-796 (2019)
- [33] Haider, W., Moustafa, N., Keshk, M., Fernandez, A., Choo, K. & Wahab, A. FGMC-HADS: Fuzzy Gaussian mixture-based coreentropy models for detecting zero-day attacks from linux systems. *Computers and Security*. **96** pp. 101906 (2020)
- [34] Moustafa, N. A new distributed architecture for evaluating AI-based security systems at the edge: Network *TONIoT* datasets. *Sustainable Cities And Society*. **72** pp. 102994 (2021)
- [35] Ogundokun, R., Awotunde, J., Misra, S., Abikoye, O. & Folarin, O. Application of machine learning for ransomware detection in IoT devices. *Artificial Intelligence For Cyber Security: Methods, Issues And Possible Horizons Or Opportunities*. pp. 393-420 (2021)
- [36] Cui, F. Deployment and integration of smart sensors with IoT devices detecting fire disasters in huge forest environment. *Computer Communications*. **150** pp. 818-827 (2020)
- [37] Benke, K. & Benke, G. Artificial intelligence and big data in public health. *International Journal Of Environmental Research And Public Health*. **15**, 2796 (2018)
- [38] Ayo, F., Folorunso, S., Abayomi-Alli, A., Adekunle, A. & Awotunde, J. Network intrusion detection based on deep learning model optimized with rule-based hybrid feature selection. *Information Security Journal: A Global Perspective*. **29**, 267-283 (2020)
- [39] Mohammadi, M., Al-Fuqaha, A., Sorour, S. and Guizani, M. Deep learning for IoT big data and streaming analytics: A survey. *IEEE Communications Surveys and Tutorials*. **20**, 2923-2960 (2018)
- [40] Singh, S., Rathore, S. & Park, J. Blockiotintelligence: A blockchain-enabled intelligent IoT architecture with artificial intelligence. *Future Generation Computer Systems*. **110** pp. 721-743 (2020)
- [41] Atitallah, S., Driss, M., Boulila, W. & Ghézala, H. Leveraging Deep Learning and IoT big data analytics to support the smart cities development: Review and future directions. *Computer Science Review*. **38** pp. 100303 (2020)
- [42] Dorri, A., Kanhere, S. & Jurdak, R. Towards an optimized blockchain for IoT. *Proceedings Of The Second International Conference On Internet-of-Things Design And Implementation*. pp. 173-178 (2017)
- [43] Panarello, A., Tapas, N., Merlino, G., Longo, F. & Puliafito, A. Blockchain and iot integration: A systematic survey. *Sensors*. **18**, 2575 (2018)
- [44] Kim, M., Hilton, B., Burks, Z. & Reyes, J. Integrating blockchain, smart contract-tokens, and IoT to design a food traceability solution. *2018 IEEE 9th Annual Information Technology, Electronics And Mobile Communication Conference (IEMCON)*. pp. 335-340 (2018)
- [45] Sharma, A., Podoplelova, E., Shapovalov, G., Tselykh, A. & Tselykh, A. Sustainable smart cities: convergence of artificial intelligence and blockchain. *Sustainability*. **13**, 13076 (2021)
- [46] Balakrishna, C. Enabling technologies for smart city services and applications. *2012 Sixth International Conference On Next Generation Mobile Applications, Services And Technologies*. pp. 223-227 (2012)
- [47] Obaidat, M. & Nicopolitidis, P. Smart cities and homes: Key enabling technologies. (Morgan Kaufmann,2016)
- [48] Trencher, G. Towards the smart city 2.0: Empirical evidence of using smartness as a tool for tackling social challenges. *Technological Forecasting And Social Change*. **142** pp. 117-128 (2019)
- [49] Chib, A., Alvarez, K. & Todorovic, T. Critical Perspectives on the Smart City: Efficiency Objectives vs Inclusion Ideals. *Journal Of Urban Technology*. **29**, 83-99 (2022)
- [50] SwarnaSudha Detecting Website Defacement using Machine Learning Techniques. *Journal of Next Generation Technology* .

- JNxtGenTech.* **1**, 47-55 (2021)
- [51] Hassan, R., Zeebaree, S., Ameen, S., Kak, S., Sadeeq, M., Ageed, Z., AL-Zebari, A. & Salih, A. State of art survey for iot effects on smart city technology: challenges, opportunities, and solutions. *Asian Journal Of Research In Computer Science.* **8**, 32-48 (2021)
- [52] Lohokare, J., Dani, R., Rajurkar, A. and Apte, A. An IoT ecosystem for the implementation of scalable wireless home automation systems at smart city level. *TENCON 2017-2017 IEEE Region 10 Conference.* pp. 1503-1508 (2017)
- [53] Grycan, W. Legislative support for improving sustainable and smart electricity consumption in polish residential sector. *Journal Of Cleaner Production.* **266** pp. 121995 (2020)
- [54] Abiodun, M., Adeniyi, E., Awotunde, J., Bhoi, A., AbdulRaheem, M. and Oladipo, I. A framework for the actualization of green cloud-based design for smart cities. *IoT And IoE Driven Smart Cities.* pp. 163-182 (2021)
- [55] Oladipo, I., AbdulRaheem, M., Awotunde, J., Bhoi, A., Adeniyi, E. & Abiodun, M. Machine learning and deep learning algorithms for smart cities: a start-of-the-art review. *IoT And IoE Driven Smart Cities.* pp. 143-162 (2021)
- [56] Allam, Z. & Jones, D. Future (post-COVID) digital, smart and sustainable cities in the wake of 6G: Digital twins, immersive realities and new urban economies. *Land Use Policy.* **101** pp. 105201 (2021)
- [57] AbdulRaheem, M., Balogun, G., Abiodun, M., Taofeek-Ibrahim, F., Tomori, A., Oladipo, I. & Awotunde, J. An enhanced lightweight speck system for cloud-based smart healthcare. *Applied Informatics: Fourth International Conference, ICAI 2021, Buenos Aires, Argentina, October 28-30, 2021, Proceedings 4.* pp. 363-376 (2021)
- [58] Dang, L., Kyeong, S., Li, Y., Wang, H., Nguyen, T. and Moon, H. Deep learning-based sewer defect classification for highly imbalanced dataset. *Computers and Industrial Engineering.* **161** pp. 107630 (2021)
- [59] Chen, D., Wawrzynski, P. & Lv, Z. Cyber security in smart cities: a review of deep learning-based applications and case studies. *Sustainable Cities And Society.* **66** pp. 102655 (2021)
- [60] Jaw, E. & Wang, X. Feature selection and ensemble-based intrusion detection system: an efficient and comprehensive approach. *Symmetry.* **13**, 1764 (2021)
- [61] Liang, Y., Li, H., Guo, B., Yu, Z., Zheng, X., Samtani, S. & Zeng, D. Fusion of heterogeneous attention mechanisms in multi-view convolutional neural network for text classification. *Information Sciences.* **548** pp. 295-312 (2021)
- [62] Mokhtari, S., Abbaspour, A., Yen, K. & Sargolzaei, A. A machine learning approach for anomaly detection in industrial control systems based on measurement data. *Electronics.* **10**, 407 (2021)
- [63] Quincozes, S., Mossé, D., Passos, D., Albuquerque, C., Ochi, L. & Santos, V. On the performance of GRASP-based feature selection for CPS intrusion detection. *IEEE Transactions On Network And Service Management.* **19**, 614-626 (2021)
- [64] Limon-Cantu, D. & Alarcon-Aquino, V. Multiresolution dendritic cell algorithm for network anomaly detection. *PeerJ Computer Science.* **7** pp. e749 (2021)
- [65] Yue, Y., Li, S., Legg, P. & Li, F. Deep learning-based security behaviour analysis in IoT environments: A survey. *Security And Communication Networks.* **2021** pp. 1-13 (2021)
- [66] Chen, J., Tian, Z., Cui, X., Yin, L. & Wang, X. Trust architecture and reputation evaluation for internet of things. *Journal Of Ambient Intelligence And Humanized Computing.* **10** pp. 3099-3107 (2019)
- [67] Malik, S., Dedeoglu, V., Kanhere, S. & Jurdak, R. Trustchain: Trust management in blockchain and iot supported supply chains. *2019 IEEE International Conference On Blockchain (Blockchain).* pp. 184-193 (2019)
- [68] Chai, H., Leng, S., Zhang, K. & Mao, S. Proof-of-reputation based-consortium blockchain for trust resource sharing in internet of vehicles. *IEEE Access.* **7** pp. 175744-175757 (2019)
- [69] Zhuang, Q., Liu, Y., Chen, L. & Ai, Z. Proof of reputation: A reputation-based consensus protocol for blockchain based systems. *Proceedings Of The 1st International Electronics Communication Conference.* pp. 131-138 (2019)
- [70] Calvaresi, D., Mattioli, V., Dubovitskaya, A., Dragoni, A. & Schumacher, M. Reputation management in multi-agent systems using permissioned blockchain technology. *2018 IEEE/WIC/ACM International Conference On Web Intelligence (WI).* pp. 719-725 (2018)
- [71] Alkadi, O., Moustafa, N., Turnbull, B. & Choo, K. A deep blockchain framework-enabled collaborative intrusion detection for protecting IoT and cloud networks. *IEEE Internet Of Things Journal.* **8**, 9463-9472 (2020)
- [72] Kolias, C., Kambourakis, G., Stavrou, A. and Gritzalis, S. Intrusion detection in 802.11 networks: Empirical evaluation of threats and a public dataset. *IEEE Communications Surveys and Tutorials.* **18**, 184-208 (2015)
- [73] Hamza, A., Gharakheili, H., Benson, T. & Sivaraman, V. Detecting volumetric attacks on iot devices via sdn-based monitoring of mud activity. *Proceedings Of The 2019 ACM Symposium On SDN Research.* pp. 36-48 (2019)
- [74] Naeem, H., Guo, B., Naeem, M., Ullah, F., Aldabbas, H. and Javed, M. Identification of malicious code variants based on image visualization. *Computers and Electrical Engineering.* **76** pp. 225-237 (2019)
- [75] Kalash, M., Rochan, M., Mohammed, N., Bruce, N., Wang, Y. & Iqbal, F. Malware classification with deep convolutional neural networks. *2018 9th IFIP International Conference On New Technologies, Mobility And Security (NTMS).* pp. 1-5 (2018)
- [76] Kumar, R., Xiaosong, Z., Khan, R., Ahad, I. & Kumar, J. Malicious Code Detection Based on Image Processing Using Deep Learning. *Proceedings Of The 2018 International Conference On Computing And Artificial Intelligence.* pp. 81-85 (2018)
- [77] Alharbi, A., Alosaimi, W., Alyami, H., Rauf, H. & Damaševičius, R. Botnet attack detection using local global best bat algorithm for industrial internet of things. *Electronics.* **10**, 1341 (2021)
- [78] Azeez, N., Odufuwa, O., Misra, S., Oluranti, J. & Damaševičius, R. Windows PE malware detection using ensemble learning. *Informatics.* **8**, 10 (2021)
- [79] Toldinas, J., Venčkauskas, A., Damaševičius, R., Grigaliūnas, Š., Morkevičius, N. & Baranauskas, E. A novel approach for network intrusion detection using multistage deep learning image recognition. *Electronics.* **10**, 1854 (2021)
- [80] Eden, S. The work of environmental governance networks: Traceability, credibility and certification by the Forest Stewardship

- Council. *Geoforum*. **40**, 383-394 (2009)
- [81] Abikoye, O., Ojo, U., Awotunde, J. & Ogundokun, R. A safe and secured iris template using steganography and cryptography. *Multimedia Tools And Applications*. **79** pp. 23483-23506 (2020)
- [82] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V. & Others Scikit-learn: Machine learning in Python. *The Journal Of Machine Learning Research*. **12** pp. 2825-2830 (2011)
- [83] Stein, G., Chen, B., Wu, A. & Hua, K. Decision tree classifier for network intrusion detection with GA-based feature selection. *Proceedings Of The 43rd Annual Southeast Regional Conference-Volume 2*. pp. 136-141 (2005)
- [84] Goswami, S., Das, A., Chakrabarti, A. & Chakraborty, B. A feature cluster taxonomy based feature selection technique. *Expert Systems With Applications*. **79** pp. 76-89 (2017)
- [85] Aslahi-Shahri, B., Rahmani, R., Chizari, M., Maralani, A., Eslami, M., Golkar, M. & Ebrahimi, A. A hybrid method consisting of GA and SVM for intrusion detection system. *Neural Computing And Applications*. **27** pp. 1669-1676 (2016)
- [86] Hsu, C., Hsieh, H., Prakosa, S., Azhari, M. & Leu, J. Using long-short-term memory based convolutional neural networks for network intrusion detection. *Wireless Internet: 11th EAI International Conference, WiCON 2018, Taipei, Taiwan, October 15-16, 2018, Proceedings 11*. pp. 86-94 (2019)
- [87] Lin, C., He, D., Zeadally, S., Kumar, N. & Choo, K. SecBCS: a secure and privacy-preserving blockchain-based crowdsourcing system. *Science China Information Sciences*. **63** pp. 1-14 (2020)
- [88] Kochovski, P., Gec, S., Stankovski, V., Bajec, M. & Drobintsev, P. Trust management in a blockchain based fog computing platform with trustless smart oracles. *Future Generation Computer Systems*. **101** pp. 747-759 (2019)
- [89] Kumar, R. & Tripathi, R. Blockchain-based framework for data storage in peer-to-peer scheme using interplanetary file system. *Handbook Of Research On Blockchain Technology*. pp. 35-59 (2020)
- [90] Keshk, M., Sitnikova, E., Moustafa, N., Hu, J. & Khalil, I. An integrated framework for privacy-preserving based anomaly detection for cyber-physical systems. *IEEE Transactions On Sustainable Computing*. **6**, 66-79 (2019)
- [91] Keshk, M., Moustafa, N., Sitnikova, E. & Benjamin Turnbull Privacy-preserving big data analytics for cyber-physical systems. *Wireless Networks*. **28** pp. 1241-1249 (2018)
- [92] Soe, Y., Feng, Y., Santosa, P., Hartanto, R. & Sakurai, K. Towards a lightweight detection system for cyber attacks in the IoT environment using corresponding features. *Electronics*. **9**, 144 (2020)
- [93] Kumar, P., Gupta, G. and Tripathi, R. TP2SF: A Trustworthy Privacy-Preserving Secured Framework for sustainable smart cities by leveraging blockchain and machine learning. *Journal Of Systems Architecture*. **115** pp. 101954 (2021)
- [94] Bhattacharya, S., Maddikunta, P., Kaluri, R., Singh, S., Gadekallu, T., Alazab, M. and Tariq, U. A novel PCA-firefly based XGBoost classification model for intrusion detection in networks using GPU. *Electronics*. **9**, 219 (2020)
- [95] Shafiq, M., Tian, Z., Sun, Y., Du, X. and Guizani, M. Selection of effective machine learning algorithm and Bot-IoT attacks traffic identification for internet of things in smart city. *Future Generation Computer Systems*. **107** pp. 433-442 (2020)
- [96] Shafiq, M., Tian, Z., Bashir, A., Du, X. and Guizani, M. IoT malicious traffic identification using wrapper-based feature selection mechanisms. *Computers and Security*. **94** pp. 101863 (2020)

*Edited by:* Kumar Abhishek

*Special Issue on:* Machine Learning and Block-chain based Solution for Privacy and Access Control in IoT

*Received:* May 11, 2023

*Accepted:* Aug 18, 2023





## THE EFFECT OF ONLINE AND OFFLINE SPORTS SAFETY EDUCATION COMBINED WITH MOOC PLATFORMS IN PHYSICAL EDUCATION TEACHING IN COLLEGES AND UNIVERSITIES

**Abstract.** In light of Internet+, how to make network technology better serve the educational cause needs more exploration. The online and offline hybrid education model that integrates MOOC is a new attempt. The sports safety of college students is the premise for the smooth development of sports activities. Therefore, a mixed teaching mode of sports safety combined with MOOC is designed to evaluate the teaching effect. However, under this teaching mode, the commonly used teaching effect evaluation methods cannot adhere to formative evaluation standards. Consequently, to better evaluate the MOOC teaching mode, a model for evaluating instructional effects based on RF mixed teaching mode is constructed. Aiming at the defects of RF in data processing, a genetic algorithm and particle swarm algorithm are used to optimize random forest. The outcomes demonstrate that the enhanced PSO-RF evaluation model has a 98.68% accuracy rate, which is 5.44% and 3.49% higher than the RF and GA-RF model respectively. Therefore, the enhanced PSO-RF-based teaching effect assessment model can better assess the mixed teaching mode in sports safety, meeting the evaluation requirements for students' learning effects.

**Key words:** MOOC; Sports safety; Physical education teaching in colleges and universities; Random Forest

**1. Introduction.** With the popularization of national fitness awareness, various sports have gradually received widespread attention. Effective physical exercise is an important guarantee for improving the body's constitution and quality of life. As the main force for the future development of society, college students should integrate physical exercise into their daily learning and life. As an important participant in sports activities, college students are also the main group of sports safety accidents. The frequent sports safety accidents in universities not only affect students' enthusiasm for participating in physical exercise, but also have adverse consequences for physical education and daily teaching activities in universities. Sports safety is the prerequisite and foundation for ensuring students' participation in sports activities [3, 4, 5]. However, most of the existing sports safety education is taught orally by teachers before physical exercise, lacking awareness of sports injuries and the ability to handle accidents. Therefore, carrying out sports safety education for college students and improving their safety awareness plays an important role. With the advancement of online education technology, sports safety education in universities has also taken on new development forms. To improve the quality and effectiveness of sports safety teaching in universities, ensure that students have the ability to handle sudden sports safety accidents, and reduce injuries, an online and offline mixed teaching (OAOMT) mode of sports safety education based on the MOOC platform is proposed. Aiming at the sports safety teaching effect under the mixed mode, the evaluation model of sports safety teaching effect based on Random Forest (RF) algorithm is constructed. Then the Particle Swarm Optimization (PSO) is used to improve the evaluation model of Random Forest teaching effect. The teaching effect evaluation model of random forest based on particle swarm optimization (PSO-RF) algorithm is constructed. It is hoped that the problems existing in sports safety education can be corrected in time. The students' sports safety awareness can be improved, and the occurrence of injury accidents in physical exercise can be reduced.

**2. Related works.** With the frequent occurrence of sports safety accidents, sports safety has received more attention. Teenagers and children may experience cardiac abnormalities or sudden cardiac arrest caused by chest impact. Therefore, Bogue KA et al. proposed that Cardiopulmonary resuscitation training should be carried out to improve the ability of sports personnel to deal with such diseases in an emergency [6]. Jani et al. investigated the quality standards for sports safety and school management. The random sampling method was used to investigate the safety of school sports in some rural and urban areas. When they take action and follow standard operating procedures, good sport management and teacher practice can be ensured [7]. Sun C and others proposed to use Big data and intelligent technology to conduct all-round monitoring and real-time

alarm on relevant elements of sports events in view of the injuries that are easy to occur in sports events. The safety of athletes in sports events is ensured [8]. Mai et al. evaluated the safety performance of sports infrastructure based on the level and utilization status of school sports infrastructure [9]. Brown J C et al. believed that conducting curriculum education can reduce the risk of athlete injury. According to the results of the questionnaire survey, the knowledge acquisition rate of participants in the course is relatively low, making it difficult to achieve the goal of reducing athlete injuries effectively [10].

The RF algorithm is suitable for data clustering, data anomaly detection and data pivoting in various environments. The integration technique of artificial intelligence teaching materials is developed to address the difficulties of long running time and low precision. Yang proposed a method for integrating artificial intelligence teaching resources based on behavioral data analysis. The random forest algorithm classifies them and adds rewards and punishments to achieve AI integration of intelligent teaching resources. The findings indicate that this strategy can increase the overall efficiency and accuracy of the method [11]. In view of the impact of evaluation units and non-landslide sample selection methods on landslide susceptibility prediction, Shu H E established a landslide susceptibility evaluation model based on Random Forest tree Bagging classifier. RF and self-organizing feature map network RF are discussed. The results showed that the self-organizing feature map network RF has a high prediction rate and success rate [12]. To improve building energy efficiency, Liu Y et al. proposed an envelope design-based building energy consumption forecast approach. The RF model is used to estimate building energy consumption. The importance of each parameter is ranked, and then the Pearson function is applied to assess the corresponding connection. The findings indicate that the RF model has a significant advantage in building energy consumption prediction [13]. Liang applied the RF algorithm to rural revitalization. The RF method is used in data mining of art services. The significance of the transformation from independent variable to dependent variable in exploring art service strategies in rural revitalization has been determined [14]. Xu et al. used RF algorithm to evaluate the spatiotemporal dynamic characteristics of urban surface thermal environment. Based on the evaluation results, strategies for urban structural adjustment and reasonable layout were proposed to achieve urban structural optimization [15].

To sum up, the relevant research on sports safety is relatively rich, covering the safety of school sports facilities and safety issues in sports events. The relevant research content of the RF algorithm involves many fields, which has numerous applications. However, existing research shows that although sports safety education is being implemented, the actual teaching effectiveness and knowledge acquisition rate are very low. Based on this, firstly, a mixed online and offline sports safety teaching model is constructed. Then, the random forest model is constructed to evaluate the teaching effect of the teaching mode. It is expected to find the problems and deficiencies in sports safety teaching, and improve the acquisition rate of sports safety knowledge and teaching quality.

### **3. Construction of sports safety teaching effect evaluation model based on improved RF algorithm.**

**3.1. The teaching mode construction of sports safety in colleges and universities combined with MOOC.** With the support of network technology, the rapid development of online education based on the MOOC platform enables modern education to break through the limitations of time and place. Physical education has the uniqueness in terms of teaching format and content. In the actual teaching process, more attention is paid to the teaching of motor skills. However, there is relatively little learning about related content, including theoretical knowledge of sports, prevention and handling of sports safety accidents, etc [16]. For the teaching of sports safety awareness, the advantages of MOOC platform is used to build a hybrid education mode for sports safety accidents. Figure 3.1 depicts the specific process.

Specifically, the OAOMT mode is mainly classified into three parts. The first part is the online learning stage before class, including discuss tasks, exercises and assessment content, etc. Students must finish the necessary theoretical knowledge study within the time-frame indicated, and master the related injury treatment skills. The second stage is offline teaching, mainly focusing on physical exercise. Assisted by reviewing relevant theoretical knowledge and answering questions, it helps students improve their safety awareness. In face-to-face communication with students, physical education teachers help students answer questions left over from the previous stage. Some additional knowledge is supplemented to help students consolidate and summarize their skills. The third stage is the application stage of knowledge and skills, which belongs to the stage of

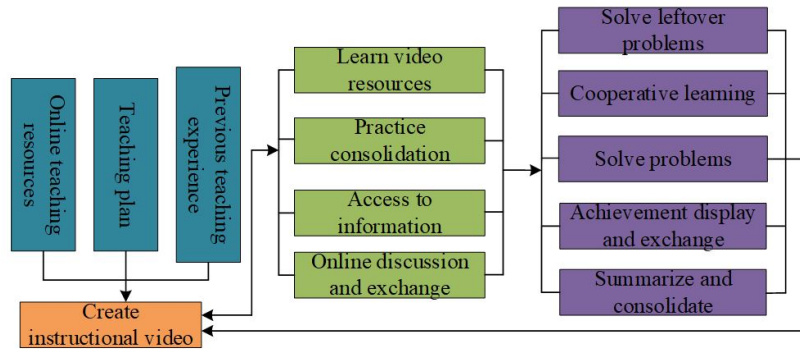


Fig. 3.1: Sports safety teaching mode under mixed mode

long-term benefit. After online and offline learning, students can simulate the injury to consolidate and output what they have learned. In the later stage of physical exercise, safety knowledge in sports should be kept in mind. In a sports injury, the injury can be dealt with urgently to minimize the extent of the injury. The online and offline hybrid teaching mode makes up for the shortcomings of offline physical education teaching. However, this teaching approach places greater demands on both students and teachers. Students need to finish pre-class learning tasks independently and master relevant basic knowledge. The handling of injuries can be proficiently mastered. At the same time, the video content must be learned before class. By publishing appropriate discussion and evaluation assignments, students' mastery of online basic knowledge is checked to ensure the normal teaching process. Relying on the MOOC platform, this mixed teaching mode realizes the teaching process of students' independent learning, discussion and evaluation. Students really become the main participants in the classroom, actively dig for knowledge, build the connection between knowledge, and effectively change traditional classroom teaching. Students receive more participation space. Under this teaching mode, the assessment of educational impact needs to adopt multiple evaluation ways, including the formative evaluation and summative evaluation, to evaluate students' learning effect objectively.

**3.2. Construction of assessment model of sports safety teaching effect based on RF algorithm.**

RF algorithm is a widely employed machine learning algorithm, which has a positive impact on data classification, regression and other processing. It is mainly consisting of a combination of Bootstrap aggregating (Bagging) and Classification and regression tree (CART) algorithms. The essential building block for the multi-style combination is the decision tree (DT), which uses the voting method to classify the data [17]. When using the RF algorithm for data classification, the Bagging method is used to randomly generate a DT set. Then the best classification features are determined by evaluation indicators such as information gain. The Bootstrap sampling method is used for data classification. A dataset of size  $D$  is extracted from dataset  $n$  (the total sample size is  $N$ ) and replace it with  $K$  to form a sample set. The probability of each sample being drawn in the training set is shown in Equation 3.1.

$$p = \left(\frac{1}{N}\right)^N \tag{3.1}$$

According to the Equation 3.1, the classifier model is constructed using the classification regression tree CART method. A KDT is trained separately. After training the sample set according to CART method, DTs are combined into RF model  $\{g_i, i = 1, 2, ..k\}$ . The samples  $q$  to is used to test the model. The test results are shown in Equation 3.2.

$$Q = \{g_1(q), g_2(q)....g_k(q)\} \tag{3.2}$$

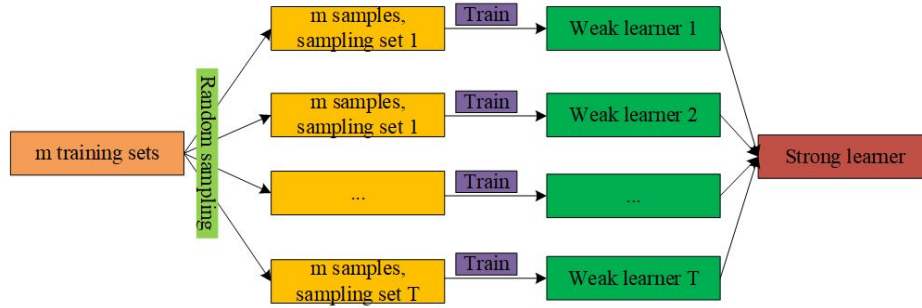


Fig. 3.2: Schematic Diagram of Bagging Method

After the classification results are obtained, the results of each DT are counted. The final regression result is the average value of all the forecast findings, as shown in Equation 3.3.

$$T(q) = \frac{1}{K} \sum_{i=1}^K t_i(q) \tag{3.3}$$

In Equation 3.3,  $q$  represents the sample data.  $t_i(q)$  represents the classification and statistical results of each sample. When the Bagging algorithm regenerates the DT set, it aggregates the operation results after randomly selecting a subset of the original data for independent operation. In Figure 3.2 the calculating procedure is displayed. The Bagging algorithm is used to increase data selection capacity for generalization and reduce overfitting problems that may occur in decision trees (DTs). When the DT node position is split, the CART algorithm mostly chooses the attribute with the Gini index (Gini) as the split attribute [18]. The impurity of the data collection is represented by the Gini index. The definition is shown in Equation 3.4.

$$Gini(D) = 1 - \sum_{i=1}^m p_i^2 \tag{3.4}$$

In Equation 3.4,  $D$  represents the data sample.  $K$  represents the scale of the DT.  $p_i$  represents  $C_j$  the probability that the sample data in sample set belongs to  $D$ . Taking an attribute  $A$  in the sample set  $D$  as an example, the sample is based on the sum of two data subsets divided by the attribute  $D_1$  and  $D_2$ . The Gini index obtained by dividing the two subsets is the weighted sum of the impurity in subsets. Equation 3.5 displays the computation process.

$$Gini_A(D) = \frac{|D_1|}{D} Gini(D_1) + \frac{|D_2|}{D} Gini(D_2) \tag{3.5}$$

The impurity of the attribute -based subset can be obtained from Equation 3.5, as shown in Equation 3.6.

$$\Delta Gini(A) = Gini(D) - Gini_A(D) \tag{3.6}$$

When generating a regression tree, it is necessary to select the data set attribute with the smallest Gini index as the node splitting attribute [19, 20]. After completing the Bagging process and the CART algorithm, the random forest model can be constructed. The specific operation process is shown in Figure 3.3 From the figure above, for classification results of the DT, the weighted voting method is used to perform weighted calculations on the classification outcomes of each DT. The final classification result is determined. If the classification result of the sample data is  $a$ , the total number of votes belonging to the  $X$  class is  $S_a$ . Equation 3.7 shows

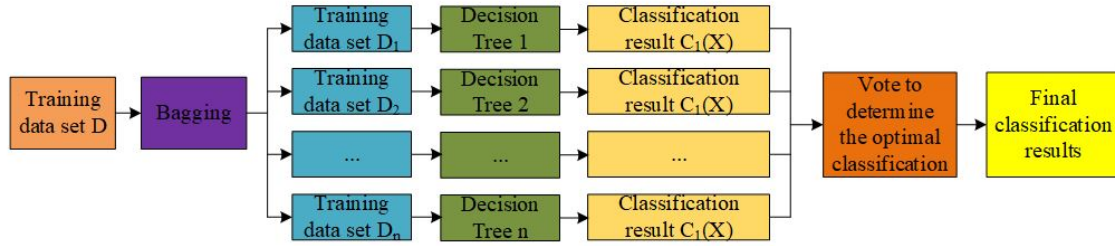


Fig. 3.3: Schematic Diagram of Random Forest

the calculation process of voting results.

$$S_a = \sum_{t=1}^K (T_{a,X})(X)W_t \tag{3.7}$$

In formula (7),  $K$  represents the scale of the DT.  $W_t$  represents the decision weight.  $T$  represents the scale of the sampling data set. If the value is 1, it means that the classification result  $T_{c,X}(X)$  of the sample  $a$  belongs to the class  $X$ . If the value is 0, it means it does not belong to the class. The generalization error of the RF is affected by correlation and strength of the DT. Generalization error has an upper limit, as shown in Equation 3.8.

$$PE^* \leq \bar{\rho}(1 - s^2)/s^2 \tag{3.8}$$

In Equation 3.8,  $\bar{\rho}$  represents the average value of the correlation between DTs. After the classification results are calculated by the weighted voting method, the category  $X$  with the highest number of votes is output as the final classification result of the sample.

**3.3. Construction of sports safety teaching effect evaluation model based on improved RF algorithm.** RF algorithm has a good effect in data classification. It needs to perform optimization calculations in parameter selection. This process takes a long time, which has a direct impact on categorization efficiency [21, 22]. Given the flaws in the execution method, the parameter values obtained by the traditional method cannot make the performance of the algorithm reach the optimum. Therefore, an intelligent approach is proposed to determine the appropriate RF technique settings. Genetic Algorithm (GA) is a non-linear optimization intelligent algorithm based on natural selection and genetic mechanism, which has good applications in the fields of parameter optimization, signal processing and machine learning [23, 24]. Firstly, the GA-RF evaluation model is constructed. After initializing the population, the crossover probability and mutation probability are determined. If  $F_{max}$  is the maximum population fitness,  $F_{avg}$  is the population average fitness, the probability of chromosome crossover  $P_c$  and mutation  $P_m$  will change with the fitness of the population. Then, the relationship between the initial probability of crossing  $P_{c1}$  and the probability of crossing  $P_{c2}$  is shown in Equation 3.9.

$$P_c = \begin{cases} P_{c1}(P_{c1} - P_{c2}) \frac{F - F_{avg}}{F_{max} - F_{avg}}, & F \geq F_{avg} \\ P_{c1}, & F < F_{avg} \end{cases} \tag{3.9}$$

Initial probability of mutation  $P_{m1}$  and the probability of mutation progress  $P_{m2}$  are shown in Equation 3.10.

$$P_m = \begin{cases} P_{m1}(P_{m1} - P_{m2}) \frac{F_{max} - F}{F_{max} - F_{avg}}, & F \geq F_{avg} \\ P_{m1}, & F < F_{avg} \end{cases} \tag{3.10}$$

Usually, mutation probabilities  $P_m$  are achieved very small. The mutation operation needs to cooperate with the crossover operation, aiming at mining the diversity of individuals in the population. Two individuals are randomly selected from the primary selection group for crossover. The obtained expression is as Equation 3.11.

$$\begin{cases} \alpha_1 = \lambda\alpha_2 + (1 - \lambda)\alpha_1 \\ \alpha_2 = \lambda\alpha_1 + (1 - \lambda)\alpha_2 \end{cases} \quad (3.11)$$

In Equation 3.11,  $\lambda \in \{0,1\}$  is any random number generated. After crossing, the chromosomes of  $\alpha_1$  and  $\alpha_2$  offspring are recalculated and optimized in the original population. In the genetic algorithm, the operators for crossover and mutation are the core of the algorithm. The crossover rate  $P_c$  and mutation rate  $P_m$  are key parameters for the convergence and stability. The  $P_c$  and  $P_m$  reflects the probability of the algorithm's crossover and mutation operations, which determines the convergence of the algorithm. Another commonly used intelligent algorithm is particle swarm optimization (PSO), which is a swarm intelligence-based random evolution method with distinct advantages in finding optimal solutions [25]. If there are particles in a population, the spatial position of each particle is  $Z$ , denoted by  $V$ , indicating the moving direction of the particle in the feasible region. The running speed of particles needs to be limited by the maximum speed and the minimum speed. The relationship between the two is obtained as shown in Equation 3.12.

$$V_{min} = -V_{max} \quad (3.12)$$

In Equation 3.12,  $V_{min}$  and  $V_{max}$  represent the particle velocity's highest and minimum values, respectively. When searching for the optimal solution, each particle will get the optimal position that each particle passes through after iterations  $p_{best}(t)$ . The solution of the minimum optimization problem is shown in Equation 3.13.

$$gbest(t) = \min\{pbest_1(t), pbest_2(t), \dots, pbest_N(t)\} \quad (3.13)$$

When the algorithm advances to the next generation, the particles in the space will update the speed and position based on previous generation information and present information. The speed update method is shown in Equation 3.14.

$$V_a^{t+1} = V_a^t + c_1 r_1 pbest_a^t - Z_a^t + c_2 r_2 (gbest_a^t - Z_a^t) \quad (3.14)$$

In Equation 3.14,  $\alpha$  represents the particle.  $c_1$  and  $c_2$  represents the learning factor of the particle.  $r_1$  and  $r_2$  are random number between 0-1.  $V_a^t$  represents the speed of the particle  $\alpha$ .  $Z_a^t$  represents the space vector  $t$  at the first iteration. The method is shown in Equation 3.15.

$$Z_a^{t+1} = Z_a^t + V_a^{t+1} \quad (3.15)$$

The new position and velocity are constantly updated through Equation 3.13 and Equation 3.14. The best positional solution is identified. Figure 3.4 depicts the specific operating process.

A larger weight coefficient is used to search the global. After obtaining the optimal solution, a smaller weight coefficient is used to search the local optimal solution to achieve the optimization effect of the particle swarm optimization algorithm.

#### 4. Performance analysis of the sports safety teaching effect assessment model based on the improved RF algorithm.

**4.1. Analysis of the training effect of the evaluation model.** To explore the teaching effect of the sports safety course in the mixed teaching mode, the teaching effect of the OAOMT mode integrated with MOOC is verified. The assessment model based on RF, the Random forest based on genetic algorithm (GA-RF) evaluation and the PSO-RF assessment model are compared and analyzed. Experimental data utilized for the research comes from the statistical data of 2500 students in the "University Physical Education" course in the initial semester of the 2017-2018 school year in a university. After preprocessing the data, 2000 pieces of data are used for model training. During the training process, the iteration times of the three models are shown in Figure 4.1.

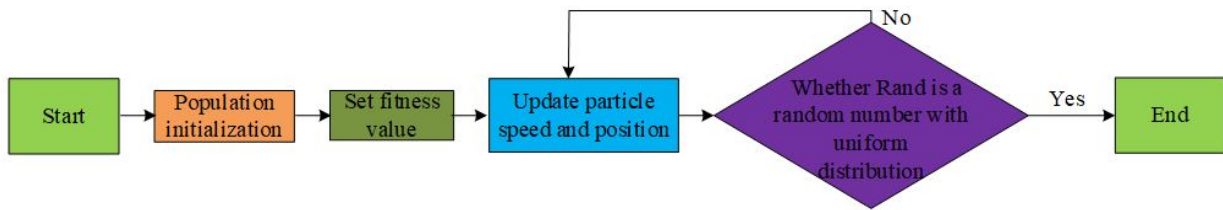


Fig. 3.4: PSO-RF algorithm flow

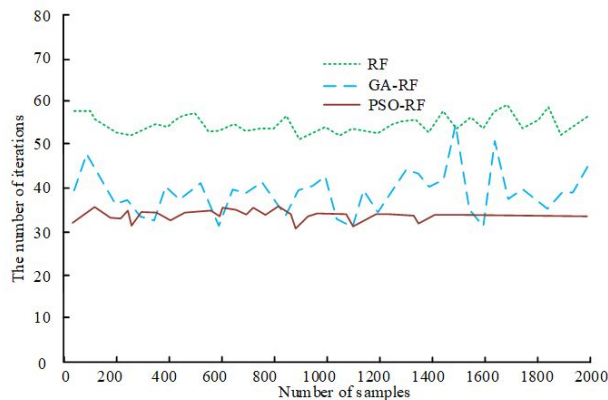


Fig. 4.1: Comparison of Iterations of Three Models

From Figure 4.1, the iteration numbers of the three models are significantly different. The RF model has undergone a significant number of iterations. The entire training process is unstable, and multiple maximum and minimum values appear. The maximum number of iterations is 60 when there are 1700 samples. The number of iterations is 52 when there are 900 samples. The average iteration is 55 times. The GA-RF model has a limited number of iterations. The rate of change is rapid, and performance stability is low. The minimal number of iterations is 32 when there are 1600 samples. The maximum iteration is 55 when there are 1500 samples. The range of iterations is 23 times, and the average number of iterations for the entire model is 42 times. The PSO-RF model has small fluctuations in the early stage of training. The number of iterations does not dramatically vary after there are 1400 samples. The average value is 34, and the convergence is good. Compared with the RF and the GA-RF, the average number of times of the PSO-RF model is 21 and 8 times lower respectively. The convergence is significantly better than the other two evaluation models. From the running results of the above model, the convergence of the improved RF model proposed in the study is significantly better than the other two methods. It has better performance in model training and can achieve convergence in fewer samples. The running time of the three models during training is shown in Figure 4.2.

From Figure 4.2, the running time of the three models diverge significantly. Specifically, the running time of the RF fluctuates greatly under different samples. The maximum running duration is 1.2s when there are 1400 samples. The minimal running time is 0.8s when there are 1500 samples. There are multiple maxima and minima throughout the test. The model performance is highly unstable. During the running process, the GA-RF model is more stable than the RF model. The running time is the smallest and the running efficiency is the best. When the number of samples is 1450, the maximum running time is 0.8s. The lowest running time is 0.56s when there are 1350 samples. The model performance is poor, with an average run time of 0.59s. The

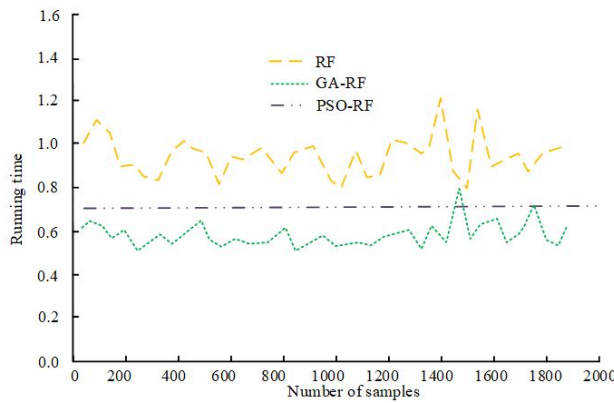


Fig. 4.2: Running time comparison of three models

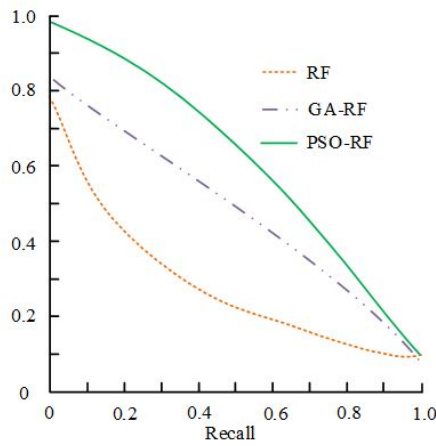


Fig. 4.3: PR comparison chart of three models

running time of the PSO-RF has no obvious fluctuation. The whole process is kept at about 0.7s. Compared with the RF model, the average running time of the PSO-RF model is 0.2s lower. Compared to the GA-RF, the average running time of the PSO-RF is slightly higher by 0.11s, but the stability of this model is completely better than that of the RF and the GA-RF.

**4.2. Analysis of the PR effect of the evaluation model.** To measure the capacity of each model, the evaluation results are evaluated by PR curves. The precision - recall (PR) curves of the three evaluation models are shown in Figure 4.3.

From Figure 4.3, when the recall rate approaches 0, the precision rate of RF is 0.78. The precision rate of GA-RF evaluation model is 0.82. The precision rate of PSO-RF model is 0.98. PSO-RF has a far greater accuracy than the other two assessment models. This indicates that the improved RF model has higher accuracy and recall rate. The performance of the model is better. Reflected in the evaluation of teaching effectiveness, the proposed model can more accurately evaluate students' learning effectiveness.

**4.3. Analysis of the application effect of the evaluation model.** The model proposed in the study is used to evaluate the actual teaching effectiveness. The actual scores of students are fitted with the evaluation scores. The results are shown in Figure 4.4.



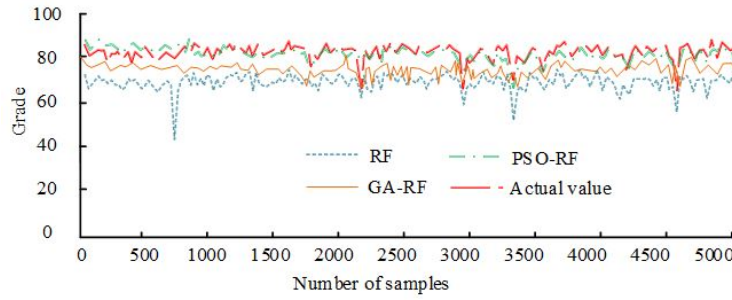


Fig. 4.4: Fitting of two models

Table 4.1: Student Performance Evaluation Results

	<i>N</i>	Min. Value	Max. Value	Mean Value	Standard Deviation
Online Learning	100	15	100	86.13	9.420
Online Discussion	100	35	100	91.24	5.093
Examination	100	50	99	92.35	7.573
Group Work	100	45	99	83.68	9.652
Classroom Performance	100	55	100	78.65	15.791
Total Score	100	58	97	82.74	7.106

In Figure 4.4 there is a large error between the student grades obtained by the RF evaluation model and the actual course grades. When the student sample is roughly between [5001000], the maximum error occurs. The score is only 45. The student achievement obtained under the model is significantly lower than the actual achievement. The extreme value of the GA-RF evaluation model changes slightly, but the grades obtained are slightly smaller than the actual grades. When the sample size of the PSO-RF model is small, the results are slightly higher than the actual results. However, as the sample size increases, the error between the evaluation results of the model and the actual results gradually decreases. From this analysis result, the optimized RF model proposed in the study has relatively small errors. During the entire evaluation period, it has a high degree of stability and accuracy, which can better reflect students' actual learning situation and achieve learning effects on sports safety courses. The performance of 5000 students who received combined physical teaching is rated. The statistical results are shown in Table 4.1.

From Table 4.1, under the OAOAMT mode, the highest score of students in online video learning of sports safety is 100 points, the lowest score is only 15 points, the average score is 86.13 points, and the standard deviation is 9.420. This indicates that students have a higher level of learning in online video courses. Online MOOC resources are more attractive to students. In the online discussion part, the highest score is 100 points, the standard deviation is 5.093, the lowest score is 35, the average is 91.24, and the lowest score is 35. The standard deviation of this learning content is the smallest and the learning level is the highest. The standard deviation of student classroom performance is 15.791, indicating a significant difference in academic performance. The polarization of students' classroom performance is evident, with some students having weaker classroom performance abilities. The standard deviation of the total student score is 7.106, indicating that there is a small difference in final grades among all students. The blended teaching mode that integrates MOOC can meet the real needs of various students. The accuracy rate of the trained evaluation model in student activism safety performance is shown Figure 4.5.

From Figure 4.5, the accuracy rates of the three models in the application process are quite different. The accuracy of the three evaluation models increases with the test sample. Among them, the RF model's accuracy

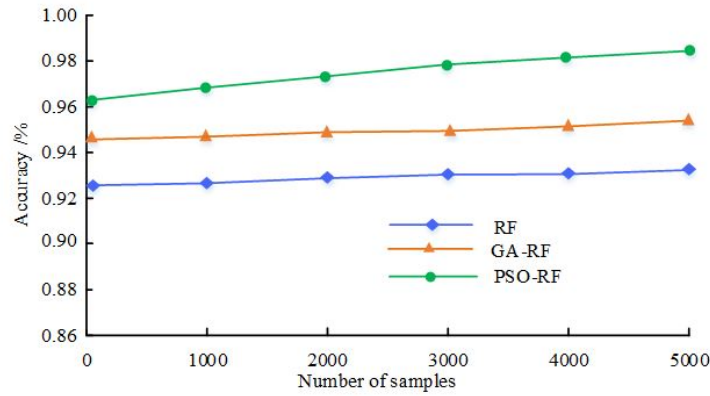


Fig. 4.5: Comparison of accuracy of three evaluation models in application

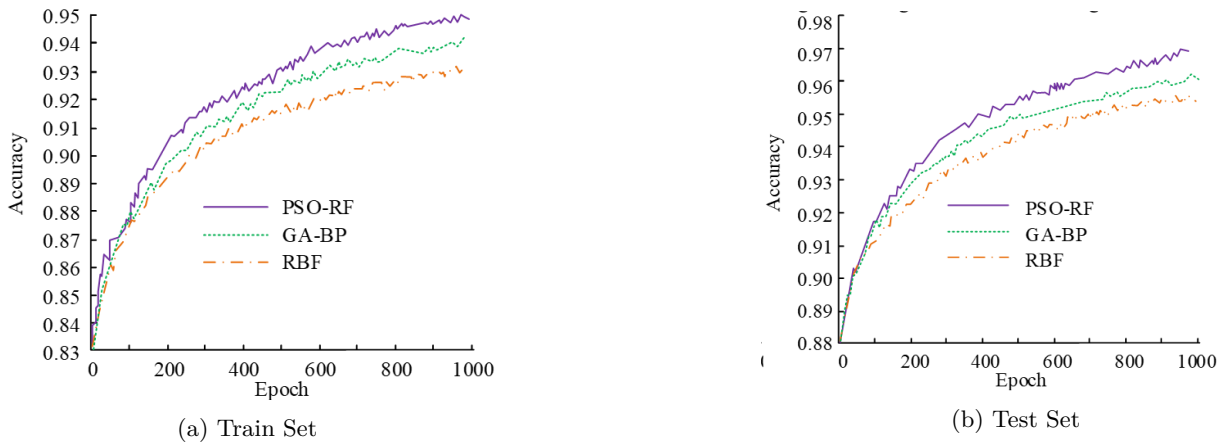


Fig. 4.6: Accuracy comparison of commonly used evaluation models

is 93.24%, and the GA-RF model is 95.19%. The PSO-RF model varies greatly. The accuracy in equilibrium state reaches 98.68%. Compared to the RF and GA-RF model, the PSO-RF model is greater 5.44% and 3.49%, respectively. The PSO-RF has a good application impact, which satisfies the assessment requirements of sports safety courses in the mixed teaching ways. It achieves a reasonable and objective evaluation of students' grades. In the intelligent evaluation of data, the commonly used K-means based radial basis function model (RBF), and genetic algorithm based optimization BP (GA-BP) are compared with PSO-RF model. The F1 values of the three models on the training and testing sets are shown in Figure 4.6. In Figure 4.6(a), on the training set, the accuracy of the PSO-RF model is 94.86%. The accuracy of the GA-BP and RBF models are 93.57% and 92.94%, respectively. In Figure 4.6(b), on the test set, the accuracy of the PSO-RF model reached 97.15%, which is higher than the GA-BP model and RBF model. This indicates that the evaluation model proposed in the study has significant advantages over similar methods.

**5. Conclusion.** Online education is in high demand due to the Internet's rapid development in the field of education. The OAOMT mode that integrates MOOC not only conforms to the development of teaching reform, but also can meet the shortcomings of sports safety education in physical education courses. Aiming at the mixed teaching mode of sports safety integrated with MOOC, an assessment methodology is constructed

for educational effects based on RF. However, the traditional RF has certain defects in the data processing process. Therefore, an assessment model based on PSPO algorithm is proposed to enhance the sports safety teaching effect assessment model. The outcomes of the trial indicate the average running time of the PSO-RF is only 0.7s. The accuracy rate reaches 0.98. the fitting degree between the assessment results and the actual results is basically consistent. The standard deviation of the total score for students is 7.106, which means that the difference in final grades for all students is very small. From the above analysis, the teaching effectiveness evaluation model based on improved RF proposed in the study has high accuracy and significantly outperforms other commonly used teaching effectiveness evaluation models. By using this evaluation method, students' grades can be objectively evaluated. Their learning situation and academic achievements in sports safety courses are well presented. From the application situation, the teaching effect of the online and offline sports safety mixed teaching mode that integrates MOOC is relatively ideal. This teaching mode can meet the actual needs of different students, which has high practical application performance. However, there are still shortcomings in the research. This model is only tested in physical education safety teaching in universities. In future research, the model should be tested in different educational and disciplinary fields to improve the applicability.

## REFERENCES

- [1] Chao, Y., Guo, X. & Wang, R. Exploration of Teaching and Education: Integration of Sports and Medicine in Physical Education in Colleges and Universities. *Journal Of Contemporary Education Research*. **6**, 63-67 (2022)
- [2] Huang, C. The Current Situation and Improvement Strategy of Physical Education in Colleges and Universities. *Journal Of Contemporary Education Research*. **5**, 1-4 (2021)
- [3] Lee, K., Won, Y. & Kim, S. The Effect of Sports Instructor's Safety Awareness and Safety Accident Prevention Activities on Sports Safety Culture. *Korean Journal Of Sports Science*. **28**, 49-59 (2019)
- [4] Slater, L., Wasserman, E. & Hart, J. Trends in Recurrent Anterior Cruciate Ligament Injuries Differ from New Anterior Cruciate Ligament Injuries in College and High School Sports: 2009-2010 Through 2016-2017: *Orthopedic Journal of Sports* 2019. (0)
- [5] Maat, S., Luppino, F., Schipper, I., Krijnen, P. & KA., B. Injury patterns after skiing and snowboarding sports accidents. *The Journal Of Sports Medicine And Physical Fitness*. **60**, 119-124 (2020)
- [6] Bogue, K., Idriss, S., Sturkey, D. & Derouin, A. Improving youth sports safety: Implementing an emergency action plan for sudden cardiac arrest. *Journal Of Pediatric Nursing*. **59**, 81-88 (2021)
- [7] Jani, M., Shahril, J., Muszali, M., Adnan, R. & Mai, S. Administrative Management Quality Standard and Teachers' Practices Compliance on School Sports Safety. *International Journal Of Academic Research In Progressive Education And Development*. **10**, 648-664 (2021)
- [8] Sun, C., Li, J. & Gao, Q. Study on the Early Warning of Athletes' Safety in Sport Events Based on the "Fusion of Data and Intelligence". *Open Journal Of Social Sciences*. **9**, 299-304 (2021)
- [9] Mai, S., Jani, J., Shahril, M., Muszali, R. & Context, B. Audit, and Index (CAI) Model Approach in Determining the Safety of Sports Infrastructure in Malaysia Schools. *IOSR Journal Of Sports And Physical Education*. **7**, 49-56 (2020)
- [10] Brown, J., Hendricks, S., Lambert, M. & Van Mw, V. BokSmart rugby safety education courses are associated with improvements in behavioral determinants in attending coaches and referees: presurvey-postsurvey study. *Injury Prevention*. **27**, 363-368 (2021)
- [11] Yang, X. Research on Integration Method of AI Teaching Resources Based on Learning Behavior Data Analysis. *International Journal Of Continuing Engineering Education And Life-Long Learning*. **30**, 492-508 (2020)
- [12] Shu, H., Abudikesimu, X., Meng, H. & Chen, K. Evaluation on landslide susceptibility based on self-organizing feature map network and random forest model: A case study of Dayu County of Jiangxi Province. *The Chinese Journal Of Geological Hazard And Control*. **33**, 132-140 (2022)
- [13] Liu, Y., Chen, H., Zhang, L. & Feng, Z. Enhancing building energy efficiency using a random forest model: A hybrid prediction approach. *Energy Reports*. **2021** pp. 7-5003 (2019)
- [14] And, L. and Solutions of Art Professional Service Rural Revitalization Strategy Based on Random Forest Algorithm. *Wireless Communications And Mobile Computing*. **2022**, 1-11 (2022)
- [15] Xu, N., Deng, F., Liu, B., Li, C., Fu, H., Yang, H. & Zhang, J. Changes in the Urban Surface Thermal Environment of a Chinese Coastal City Revealed by Downscaling MODIS LST with Random Forest Algorithm. *Journal Of Meteorological Research*. **35**, 759-774 (2021)
- [16] Kim, G. & Gurvitch, R. The effect of Sports-based Physical Activity program on teachers' relatedness, stress and exercise motivation. *Health Education Journal*. **79**, 658-670 (2020)
- [17] An, A. & Bpabc, D. Spatial landslide susceptibility assessment using machine learning techniques assisted by additional data created with generative adversarial networks. *Geoscience Frontiers*. **12**, 625-637 (2021)
- [18] Papineni, S., Reddy, A., Yarlagadda, S. & Akkinen, H. An Extensive Analytical Approach on Human Resources using Random Forest Algorithm. *International Journal Of Engineering Trends And Technology*. **69**, 119-127 (2021)
- [19] Xia, G., Chen, J., Tang, X., Zhao, L. & Sun, B. Shift quality optimization control of power shift transmission based on particle

- swarm optimization–genetic algorithm. *Proceedings Of The Institution Of Mechanical Engineers*. pp. 872-892 (2022)
- [20] Teer, B. Performance analysis of sports training based on random forest algorithm and infrared motion capture. *Journal Of Intelligent & Fuzzy Systems: Applications In Engineering And Technology*. **40**, 6853-6863 (2021)
- [21] Alatawi, M., Alsubaie, N., Ullah, K., Sadad, T., Alwageed, H., Ali, S. & Zada, I. Cyber Security against Intrusion Detection Using Ensemble-Based Approaches. *Security And Communication Networks*. **2023** pp. 2023
- [22] Arbogast Kristy B, M. Sport Safety for Adolescents: Linking Biomechanics of Repetitive Head Impacts With Health and Wellbeing. *The Journal Of Adolescent Health: Official Publication Of The Society For Adolescent Medicine*. **72**, 2023
- [23] Donnelly Peter, K. Contesting the autonomy of sport to realize the right to safe sport: a Canadian case study. *The International Sports Law Journal*. **22**, 2022
- [24] Sae Hyung Kim, H. Development of physical education safety awareness scale for middle school students using multi-facets Rasch model and group classification through latent profile analysis. *Korean Journal Of Sports Science*. **29**, 2020
- [25] Wang Yan, Z. Wei, Ge Ziyang, Wang Junliang, Xu Haoyu, Jiang Chao. Pressure sensing recognition of FBG array based on random forest algorithm. *Optoelectronics Letters*. **19**, 2023

*Edited by:* Mudasir Mohd

*Special Issue on:* Scalable Computing in Online and Blended Learning Environments: Challenges and Solutions

*Received:* May 15, 2023

*Accepted:* Jul 5, 2023



## SCALABLE COMPUTING INFRASTRUCTURE FOR ONLINE AND BLENDED LEARNING ENVIRONMENTS

**Abstract.** With the growing popularity of online learning and blended learning, as well as the rapid development of cloud computing and big data technology, scalable computing infrastructure has become an indispensable part of building a modern education platform. Method: Five experiments were conducted to test the scalability and reliability of computing infrastructure based on online and blended learning environments. The experiments include the performance comparison of online learning platforms based on different virtualization technologies, the performance comparison of online and hybrid learning environments under different loads, the comparison of online learning experiences under different bandwidth constraints, the system stability test under different user numbers, and the comparison of access speeds in different regions. Result: The experimental results showed that on an online learning platform using the KVM (Kernel-based Virtual Machine) interface, when the number of concurrent users is 99, the response time is 100.9ms, and the CPU (Central Processing Unit) utilization rate is 60.9%. Under low load conditions, the concurrent access volume is 200; the response time is 50ms, and the throughput is 10.3. When accessing locally, the latency is 9.19ms; the download speed is 500.3KB/s; the network throughput is 399.8KB/s. Conclusion: Exploring the scalability, reliability, performance, stability, and access speed of online learning platforms is crucial for improving platform competitiveness and ensuring user experience.

**Key words:** Learning Environment, Scalable Computing, Basic Facilities, Online and Blended Learning

**1. Introduction.** In the current educational environment, network technology has become an essential tool. In particular, the emergence of online and blended learning models has highlighted the need for efficient, secure and stable computing infrastructure. The traditional client server structure is no longer able to meet the needs of modern education. At the same time, a large amount of online education resources needs to be stored and managed, which also requires a strong infrastructure to support. Scalable computing infrastructure has huge advantages in improving computing efficiency and reducing costs, and has a wide range of application prospects. On the other hand, due to the large number of user visits involved in online education, it is necessary to have strong scalability. Especially during peak hours, it is necessary to be able to withstand a large amount of concurrent access. In addition, for online education, data security is also very important, so it is necessary to have a highly reliable data storage and management mechanism.

In summary, the scalable computing infrastructure of online and blended learning environments is an integral part of modern education. In order to improve the quality of online and blended learning, major educational institutions are considering using scalable computing infrastructure to support their educational needs. This technology can effectively coordinate physical resources and dynamically allocate computing power to meet the needs of educational applications. Meanwhile, due to its scalability, it can flexibly respond to educational institutions of different sizes, as well as constantly changing user numbers and access traffic.

**2. Related Work.** Massive Open Online Course (MOOC) allows learning to take place anytime and anywhere with little external monitoring by the instructor. Typically, a highly diverse group of learners participating in MOOC needs to make decisions related to their learning activities in order to achieve academic success. Therefore, supporting self-regulated learning strategies and adapting to relevant human factors such as gender, cognitive ability, and prior knowledge are considered important. Self-regulated learning has been widely studied in traditional classroom environments, but little is known about how to support self-regulated learning in MOOC. There are currently few experimental studies conducted in MOOC. To fill this gap, Wong Jacqueline conducted a systematic review of research on methods to support self-regulated learning in various types of online learning environments and how to address human factors [1]. Student engagement has increased student satisfaction, enhanced their learning motivation, reduced their sense of isolation, and improved their performance in online courses. Martin Florence examined students' perspectives on various participation

strategies used in online courses based on the Moore interaction framework. 155 students completed 38 surveys on learner-to-learner, learner-to-teacher, and learner-to-content participation strategies. The results of Martin Florence's research had implications for online teachers, instructional designers, and managers who hoped to increase their participation in online courses [2]. Blended learning creates a "rich" educational environment through a variety of technical exchanges in face-to-face and online teaching. The characteristics of students are closely related to the learning effect in the blended learning environment.

The ability of students to guide themselves in learning and utilizing learning technologies can affect their learning efficiency. Geng Shuang investigated the influence of self-directed learning, technical preparation and learning motivation of students learning subjects in blended learning and non-blended learning environments on three kinds of existence (social, teaching and cognitive) [3]. Student satisfaction is used as one of the key factors in evaluating online courses, while perceived learning is seen as an indicator of learning. Alqurashi Emtinan aimed to explore how online learning self-efficacy, learner content interaction, learner teacher interaction and learner learner interaction predict student satisfaction and perceived learning [4].

Many undergraduate classes in science and engineering courses use asynchronous computer platforms to host teaching materials, such as lecture videos or forums. These platforms also have the ability to provide students with immediate feedback on formative assessment tasks. Although many studies have been conducted on computer-based feedback, more needs to be learned about how students interact with real-time feedback and how these interactions affect their learning. Chen Xin uses the institutional version of the edX platform to describe the interaction of physics beginners through a computer-based instant simple correction feedback tool called "checkable answer function". The results indicate that certain patterns of participation in feedback reflect effective learning strategies and can significantly predict higher performance [5]. However, online and blended learning environments need more scalable and reliable computing infrastructure to support them.

Despite private schools with high-end infrastructure attempting to establish a dominant position in providing high-quality learning, it remains a distant dream for impoverished students. Internet-enabled and second-generation Internet-based blended learning environments can stimulate student engagement, motivation, and learning. Dey Priyadarshini proposed an Internet-based blended learning platform that combines traditional classroom interaction models with synchronous e-learning, with digital audiovisual content provided by professional online teachers. The results show that the hybrid learning platform in the classroom environment, coupled with high-quality digital content, professional online teachers and on-site teaching assistants as classroom focal points, can create a learning environment. Regardless of the socio-economic status of students, their academic performance and well-being can be significantly improved [6].

The rapid development of network makes people realize that blended learning is an effective learning mode. Wang used the text mining method to analyze the blended learning practice data of 17 countries provided by Christensen Institute. The factors that hindered the implementation of blended learning were analyzed by classifying and extracting textual information on the use of blended learning model selection and the challenges of blended learning. Wang Lin has provided reference for improving the efficiency of blended learning practice, especially in practice mode selection, infrastructure preparation and teacher and student ability training [7]. However, these scholars did not analyze the scalable computing infrastructure for online and blended learning environments, but only discussed it from a shallow level.

In order to solve the problem that the computing infrastructure of traditional learning platforms cannot meet the rapidly growing user needs, resulting in reduced service quality and poor user experience, this paper uses a variety of experimental methods to analyze the scalable. The experimental results showed that the online learning platform using KVM interface performed well, with a system throughput of 10.3 under low load conditions. In addition, this article also explored indicators such as network latency, download speed, and network throughput. These experimental results contributed to improving the stability and user experience of online learning platforms.

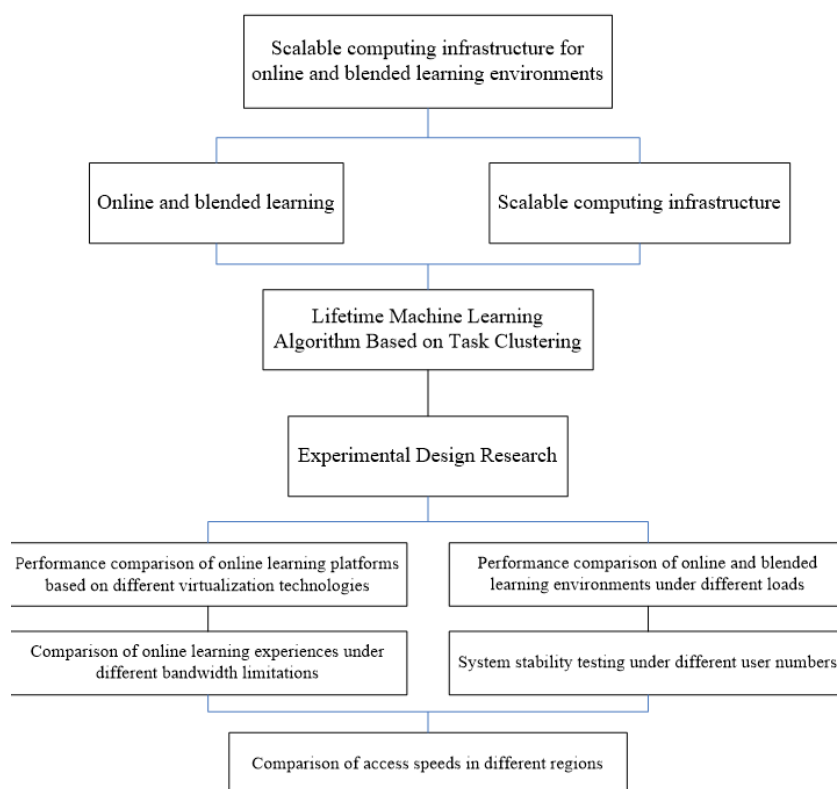


Fig. 3.1: Overall framework

### 3. Online and Blended Learning Environment.

**3.1. Online and Blended Learning.** Online learning refers to learning through the internet, where students can use computers, mobile phones, and other devices at home or anywhere to learn. The benefits of online learning are convenience, flexibility, and the ability to learn anytime and anywhere [8]. With the continuous development and upgrading of internet technology, there are increasing number of online learning resources have emerged on the internet. Blended learning refers to a teaching method that combines traditional teaching forms with online learning [9, 10]. This approach organically combines face-to-face teaching with online teaching to help students better master knowledge.

**3.2. Scalable Computing Infrastructure.** Online and blended learning need an efficient computing infrastructure to support it. This infrastructure must be reliable, secure, scalable, high-performance, and easy to manage. Scalable computing infrastructure refers to the computing resources that can be expanded according to needs to meet user needs. As the number of online and blended learning users may continue to increase or decrease, a scalable computing infrastructure is needed to quickly adapt to changes in demand [11, 12]. This paper aims to study the scalable computing infrastructure for online and blended learning environments, as shown in Figure 3.1. Implementing scalable computing infrastructure requires considering many aspects, including hardware, software, network, security, and more. Here are some methods for implementing scalable computing infrastructure:

1. **Adopting cloud computing technology:** Cloud computing technology can help achieve scalable computing infrastructure. It can provide users with elastic computing services, so that they can adjust computing resources according to their needs. In addition, cloud computing service providers can also provide high availability, load balancing, protection and other functions to ensure system stability and

security [13].

2. **Using containerization technology:** Containerization technology can help achieve scalable computing infrastructure. When using containerization technology, each application is packaged into a container, and each container is independent of each other. This means that containers can be quickly started or stopped without affecting the operation of other containers or the entire system [14].
3. **Adopting streaming data processing technology:** Streaming data processing technology can help achieve real-time data processing and analysis. By using streaming data processing technology, data can be compressed into a stream format and processed and analyzed in real-time using a streaming computing engine. Streaming data processing technology can help achieve high throughput and low latency data processing, which is very important for online and blended learning environments [15].
4. **Using real-time monitoring and automatic expansion:** Real-time monitoring and automatic scaling are important tools for implementing scalable computing infrastructure. By using real-time monitoring, problems in the system can be detected in a timely manner and corresponding measures can be taken. Automatic expansion can automatically increase or decrease computing resources based on demand to ensure that the system is always in optimal state [16].

**3.3. Lifetime Machine Learning Algorithm Based on Task Clustering.** In online and blended learning environments, algorithms need to be able to constantly learn from data, adapt to new tasks, and constantly update models [17, 18]. However, traditional machine learning algorithms are usually only suitable for specific tasks and datasets, and cannot adapt well to changing environments. In contrast, lifelong machine learning algorithms based on task clustering form a task-based learning framework by clustering similar tasks. This framework enables algorithms to respond more flexibly to new tasks while maintaining their ability to learn and predict historical data. The self-updating and iterative learning of the algorithm are realized, which makes the algorithm perform better in online and blended learning environments [19, 20]. In addition, lifelong machine learning algorithms based on task clustering can also play a role in the scalability of computing infrastructure. The lifelong machine learning algorithm based on task clustering (TC) clusters learning tasks into categories with interrelated tasks.

**3.4. Nearest neighbor generalization.** The TC algorithm classifies by nearest neighbor. To measure the similarity between data sample points, TC algorithm uses a globally weighted Euclidean distance matrix:

$$dist_f(c, u) = \sqrt{\sum_o f^{(o)} (c^{(y)} - u^{(y)})^2} \quad (3.1)$$

Here,  $f$  represents an adjustable weight influence factor vector. As a parameter of the Euclidean distance matrix,  $f$  obviously determines the generalization attribute of the nearest neighbor.

**3.5. Adjusting distance matrix.** The TC algorithm migrates knowledge between tasks by adjusting  $f$  for certain tasks, and then reuses  $f$  in the nearest neighbor generalization of other tasks. This step is achieved by minimizing the distance between training data of the same class while maximizing the distance between training data of different categories:

$$R_m(f) = \sum_{c,u} \gamma_{c,u} dist_f(c, u) \rightarrow \min \quad (3.2)$$

$f^* = arg \min_f R(f)$  represents the parameter vector obtained by minimizing  $R$ ;  $dist^*$  represents the corresponding optimal distance matrix. Through optimization,  $dist^*$  focuses on the relevant input dimensions of the math task.  $S \subset \{1, 2, \dots, M\}$  represents a subset of supported tasks:

$$f_S^* = arg \min \sum_{m \in S} R_m(f) \quad (3.3)$$

$f_S^*$  is the optimal parameter matrix of  $R$ .



**3.6. Task transfer matrix.** To measure the degree of correlation between different tasks, the TC algorithm calculates the following matrix:

$$V = (v_m, v_z) \quad (3.4)$$

Matrix  $V$  is called the task transfer matrix. The task transfer matrix contains a value  $v_{m,z}$  for task  $m$  and task  $z$ .  $v_{m,z}$  indicates the expected generalization precision of task  $m$  when the  $R$  optimal distance matrix of task  $z$  is used. Each element  $v_{m,z}$  is evaluated through cross validation of  $l$ -fold.

**3.7. Task clustering and task hierarchy.** The TC algorithm clusters  $M$  learning tasks into  $Y$  separate classes ( $Y$  is less than or equal to  $M$ ), represented by  $S_1, \dots, S_Y$ . This step is achieved by maximizing the following functions:

$$K = \frac{1}{M} \sum_{y=1}^Y \sum_{m \in S_Y} \frac{1}{|S_Y|_{z \in S_Y}} \sum_{z \in S_Y} v_{m,z} \quad (3.5)$$

$K$  measures the average estimation generalization accuracy when task  $m \in S_Y$  uses the  $R$  optimal distance matrix of task  $z \in S_Y$  in the same cluster.

When conducting machine learning on large-scale datasets, traditional single machine computing often has low efficiency. The lifelong machine learning algorithm based on task clustering supports distributed computing, dispersing computing tasks to multiple processing nodes, and improving computational efficiency. As the amount of data continues to increase, traditional batch processing methods may encounter performance bottlenecks. The lifelong machine learning algorithm based on task clustering supports streaming computing and can divide data into multiple data streams for gradual calculation and update, which avoids the overhead of data preprocessing and storage, and improves computational efficiency.

The lifelong machine learning algorithm based on task clustering utilizes already learned knowledge for rapid learning and adaptation when facing new tasks, and shares the learned knowledge among different tasks. In an online learning environment, lifelong machine learning algorithms can dynamically learn based on changes in real-time data. It can group similar tasks through clustering and perform incremental learning within each task group using previously learned models. This method can reduce repetitive learning of data, improve learning efficiency, and enable the model to better adapt to new tasks. In Blended learning environment, lifelong machine learning algorithm can combine the advantages of offline and online learning. It can extract universal features and knowledge from offline data and use them for rapid learning of new tasks. Meanwhile, online learning can help algorithms adapt to constantly changing environments and new task requirements.

**4. Infrastructure Evaluation Based on Online and Blended Learning Environment.** In order to meet the rapid growth and change of online and blended learning environments, it is necessary to establish a scalable computing infrastructure. The infrastructure should be able to support a large number of users to conduct online and blended learning at the same time, support a variety of different learning scenarios and applications, and dynamically adapt to changing needs. The main goal of this article is to test the extensibility and reliability of the computing infrastructure, that is, in the case of the increasing number of users and applications, it can still operate normally with reasonable response time and stable quality of service, and maintain high availability under various load and network conditions. This article analyzed it through five experiments.

**5. Performance Comparison of Online Learning Platforms Based on Different Virtualization Technologies.** The purpose of the experiment is to compare the performance of online learning platforms based on different virtualization technologies (such as KVM, Xen, etc.). KVM is a kernel based virtualization technology. It allows the creation of multiple virtual machines on top of the kernel of the host operating system, each of which can run its own operating system and applications. Xen is an open source virtualization platform that provides hardware virtualization and concurrency capabilities, allowing multiple independent virtual machines to run simultaneously on a physical server. Experimental content: Environmental preparation is the design and construction of online learning platform environments based on different virtualization technologies (such as KVM, Xen, etc.), and the configuration of corresponding hardware and software environments. Test case

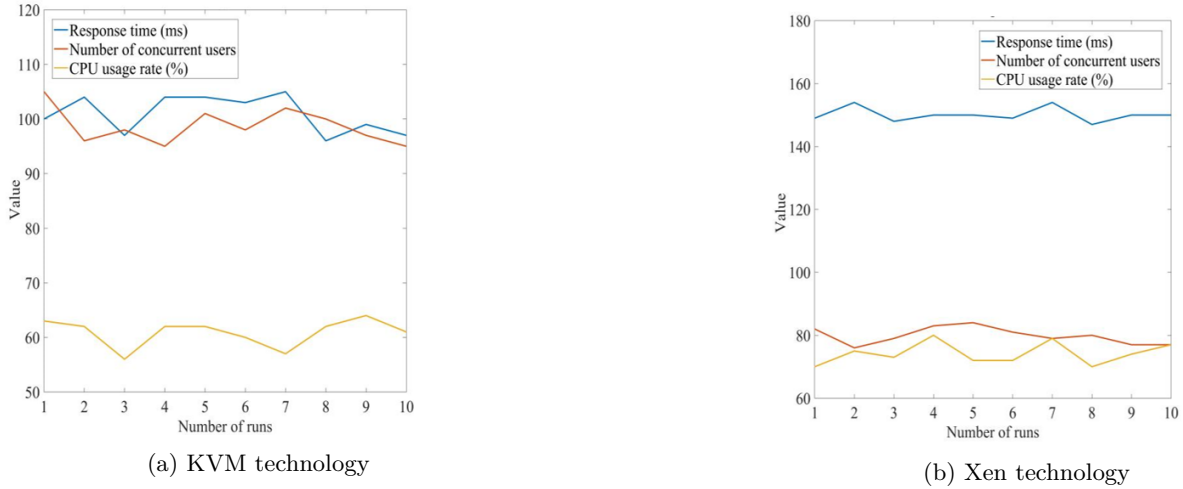


Fig. 5.1: Comparison of performance of online learning platforms based on different virtualization technologies

Table 5.1: Performance Mean Analysis

	KVM	Xen
Response time (ms)	100.9	150.1
Number of concurrent users	99	80
CPU usage rate (%)	60.9	74.2

design involves designing different types of test cases, including simulating requests from different concurrent users, simulating requests under different workloads, and so on. Running test cases is the process of running test cases under different virtualization technologies, and recording the response time, number of concurrent users, CPU usage, and other indicator data for each test case. It is necessary to run 10 times to eliminate errors and calculate the average value of each indicator. It analyzes experimental data and compares the performance of different virtualization technologies under different test cases to draw conclusions. Experimental indicator: Response time is the time required for the system to respond to requests. The number of concurrent users is the number of users accessing the system simultaneously. CPU usage is the percentage of CPU resources occupied. As shown in Figure 5.1, the performance comparison of online learning platforms based on different virtualization technologies is presented. Figure 5.1(a) shows KVM technology, and Figure 5.1(b) shows Xen technology. Table 5.1 shows the performance mean analysis.

It can be seen that on the online learning platform using KVM interface, the number of concurrent users was 99; the response time was 100.9ms; the CPU utilization rate was 60.9%. On an online learning platform using the Xen interface, the number of concurrent users was 80; the response time was 150.1ms; the CPU utilization rate was 74.2%. It indicated that although there were more concurrent users using KVM interface than those using Xen interface, the response time of online learning platforms using KVM interface was slightly faster than those using Xen interface. The percentage of CPU resources occupied by online learning platforms that simultaneously used KVM was slightly lower.

**6. Performance Comparison of Online and Blended Learning Environments under Different Loads.** The purpose of the experiment is to test the performance of online and blended learning environments under different loads. Experiment content: Environment preparation is to design and build online and blended

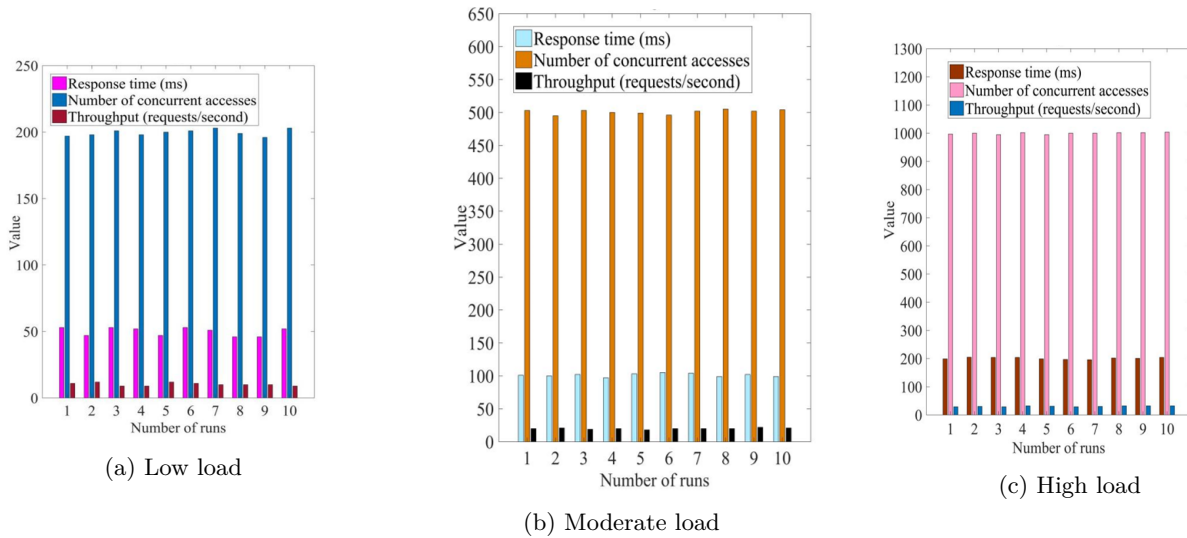


Fig. 6.1: Performance comparison of online and blended learning environments under different loads

Table 6.1: Performance mean analysis under different loads

	Low Load	Moderate Load	High Load
Response time (ms)	50	101.2	201.1
Number of concurrent accesses	200	501	1000
Throughput (requests/second)	10.3	20.1	30.6

learning environments based on scalable computing infrastructure, and configure corresponding hardware and software environments. Test case design is to design different types of test cases according to the characteristics of online and blended learning environments, such as simulating multiple users accessing the system at the same time, simulating large-scale dataset training, and setting different load values. Running test cases refers to running test cases under different loads and recording various indicator data, such as response time, concurrent access volume, throughput, etc. It is necessary to run 10 times to eliminate errors and calculate the average value of each indicator. Analyzing the experimental data is to compare the performance of online and blended learning environments under different loads, and then draw conclusions.

Experimental indicator: Response time is the time required for the system to respond to requests. Concurrent access is the number of requests that simultaneously access the system. Throughput is the number of requests processed per unit of time. Figure 6.1 shows the performance comparison of online and blended learning environments under different loads. Figure 6.1(a) shows low load; Figure 6.1(b) shows medium load; Figure 6.1(c) shows high load. Table 6.1 shows the average performance analysis under different loads.

Under low load conditions, the concurrent access volume was 200; the response time was 50ms; the throughput was 10.3. In a medium load state, the concurrent access volume was 501; the response time was 101.2ms; the throughput was 20.1. Under high load conditions, the concurrent access volume was 1000; the response time was 201.1ms; the throughput was 30.6. It indicated that as the load increased, the response time and concurrent access volume of the system would increase, and the throughput would increase. When the load was low, the system performed well and can meet user needs. When the load reached a certain level, it was necessary to adjust or expand the system to ensure its stable performance.

Table 7.1: Online Learning Experience Under Different Bandwidth Limitations

Bandwidth Limitation	Video Buffer Time (s)	Video Quality (Resolution)	Download Speed (KB/s)
Unlimited	1	1080p	400
Low bandwidth	5	480p	100
Medium bandwidth	10	360p	50
High bandwidth	1	1080p	800

**7. Comparison of Online Learning Experiences under Different Bandwidth Limitations.** The purpose of the experiment is to compare the differences in online learning experiences under different bandwidth limitations. Experimental content: Environmental preparation involves adjusting network bandwidth settings for different bandwidth limitations, building an online learning platform, and configuring corresponding hardware and software environments. Test case design is based on the characteristics of online learning platforms to design different types of test cases, such as watching videos, downloading teaching materials, and setting different bandwidth restrictions. Running test cases refers to running test cases under different bandwidth limitations and recording various indicator data, such as video buffer time, video image quality, download speed, etc. Its analysis of experimental data is to compare the online learning experience under different bandwidth constraints and draw conclusions. Experimental indicator: Video buffer time is the time spent buffering video content. Video quality refers to the quality of video playback. Download speed is the time required to download teaching materials. Table 7.1 shows the online learning experience under different bandwidth limitations. When the bandwidth limit was too high, the download speed increased, and the video buffering time and image quality were also guaranteed. When the bandwidth limit is too low, the download speed drops, the video buffer time increases, and the image quality decreases. Therefore, online learning platforms need to provide different levels of service quality or intelligently adjust bandwidth restrictions based on different user network environments to improve user experience.

**7.1. System Stability Testing under Different User Numbers.** The purpose of the experiment is to test the stability of the system under different user numbers. Experimental content: Environmental preparation involves building an online learning platform and configuring corresponding hardware and software environments. Test case design is to design different types of test cases based on the characteristics of the online learning platform, such as simulating multiple users accessing the system and performing operations at the same time, simulating different numbers of data sets for training, and setting up different numbers of users. Running test cases refers to running test cases under different user numbers and recording various indicator data, such as error rate, processing time, performance jitter, etc. It is necessary to run 10 times to eliminate errors and calculate the average value of each indicator. Analyzing experimental data is to compare the system stability under different user numbers and draw conclusions.

Experimental indicator: Error rate is the frequency of system errors. Processing time is the time required for the system to process user requests. Performance jitter is the degree of fluctuation and stability of system performance. Figure 7.1 shows the comparison of system stability tests under different user numbers. Figure 7.1(a) shows 50 users; Figure 7.1(b) shows 100 users; Figure 7.1(c) shows 200 users. Table 7.2 shows the mean analysis of system stability tests. It can be seen that when the number of users was 50, the error rate was 0.095%; the processing time was 0.1s; the performance jitter was 0.0108s. When the number of users was 100, the error rate was 0.206%; the processing time was 0.1513s; the performance jitter was 0.0155s. When the number of users was 200, the error rate was 0.504%; the processing time was 0.2s; the performance jitter was 0.02s. This indicated that while the number of users was increasing, the error rate was also increasing. Therefore, the platform needs to consider how to maintain system performance and stability while expanding user scale.

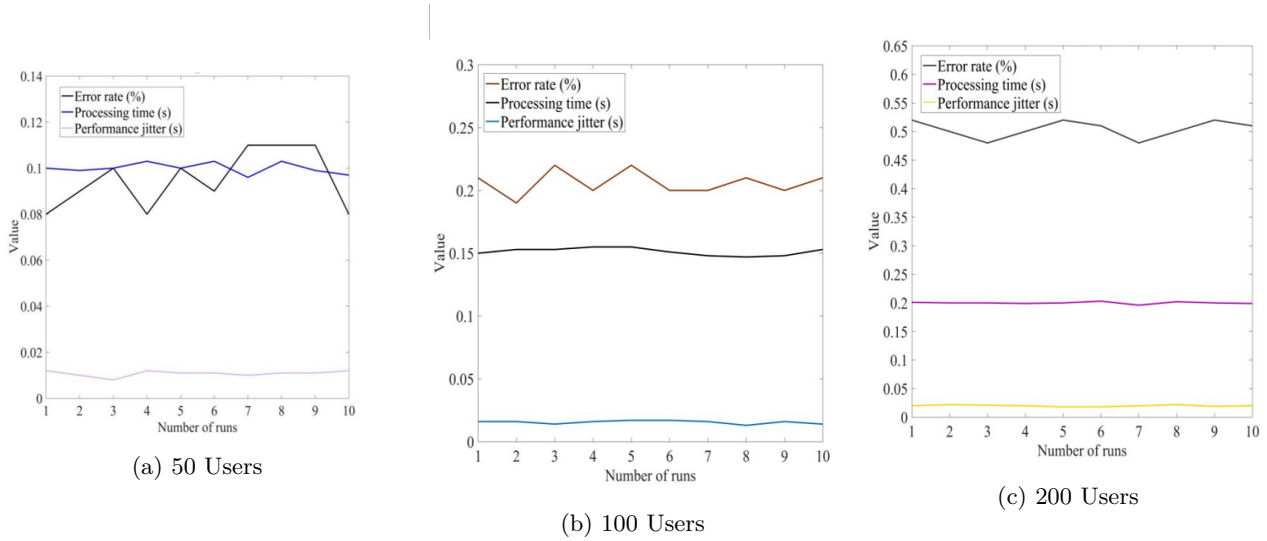


Fig. 7.1: Comparison of system stability testing under different user numbers

Table 7.2: Analysis of system stability test mean

	50 users	100 users	200 users
Error rate (%)	0.095	0.206	0.504
Processing time (s)	0.1	0.1513	0.2
Performance jitter (s)	0.0108	0.0155	0.02

**8. Comparison of Access Speed in Different Regions.** The purpose of the experiment is to compare the speed differences of accessing online learning environments in different regions. Experimental content: Environmental preparation involves setting up different network access methods for different regions, building online learning platforms, and configuring corresponding hardware and software environments. Test case design is based on the characteristics of online learning platforms to design different types of test cases, such as accessing websites, downloading teaching materials, and setting different regions. Running test cases refers to running test cases in different regions and recording various indicator data, such as latency, download speed, network throughput, etc. It is necessary to run 10 times to eliminate errors and calculate the average value of each indicator. Analyzing data is to compare the access speeds of different regions and draw conclusions.

Experimental indicator: Delay is the time required to access the system. The download speed is the time required to download teaching materials from the system. Network throughput is the amount of data transmitted by a network per unit of time. Figure 8.1 shows the comparison of access speeds in different regions, with Figure 8.1(a) showing the local area and Figure 8.1(b) showing other regions. Table 8.1 shows the analysis of mean access speed.

It can be seen that when access was made locally, the latency was 9.19ms; the download speed was 500.3KB/s; the network throughput was 399.8KB/s. When accessing in other regions, the latency was 50.7ms; the download speed was 300.8KB/s; the network throughput was 199.8KB/s. This indicates significant differences between regions, with local access speeds being faster than other regions. Therefore, when building an online learning platform, it is necessary to consider the location of server deployment and optimize network structure to improve the network access experience of users in different regions.

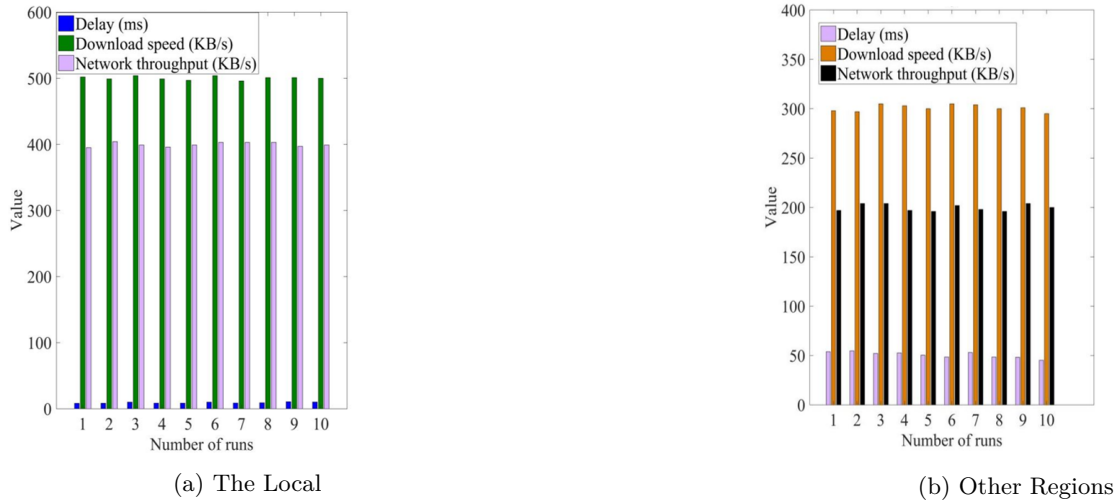


Fig. 8.1: Comparison of access speed in different regions

Table 8.1: Analysis of Average Access Speed

	This Locality	Other Regions
Delay (ms)	9.19	50.7
Download Speed (KB/s)	500.3	300.8
Network Throughput (KB/s)	399.8	199.8

**9. Conclusions.** With the gradual popularization of online education, the development of online learning platforms is receiving increasing attention from people. How to ensure the scalability, reliability, performance, stability, and access speed optimization of online learning platforms has become a hot topic in the industry. Through the experimental analysis of online and blended learning environments, the following conclusions can be drawn: the virtualization technology of online learning platform has an important impact on its performance. Under the same number of concurrent users, the response time of online learning platforms using KVM interface is slightly faster than those using Xen interface, and the percentage of CPU resources occupied is slightly lower. When the load is low, the system performs well and can meet user needs. When the load reaches a certain level, it is necessary to adjust or expand the system to ensure its stable performance. The online learning experience under different bandwidth limitations can also affect user experience, and the system needs to provide different service quality levels for intelligent adjustment based on the different network environments of users. The stability of the system under different user numbers also needs to be considered. As the number of users increases, the system may crash or run slowly. Therefore, system expansion or performance optimization is needed to ensure system stability. The farther the user is from the online learning platform server, the slower the access speed. Therefore, online learning platforms should consider using content distribution networks or multiple data centers to improve the user experience and access speed. Through the use of appropriate virtualization technology, load balancing, bandwidth control, system expansion and geographic distribution, the efficiency and stability of the online learning platform can be effectively improved, and high-quality user experience and good service performance can be achieved. However, due to the limitations of time and technology, the problems encountered in the research of online and blended learning environment in this paper have not been studied in detail, which will be further discussed in the future.

## REFERENCES

- [1] Jacqueline, W., Baars, M. & Davis, D. Tim Van Der Zee. *Geert-Jan Houben, Fred Paas. "Supporting Self-regulated Learning In Online Learning Environments And MOOCs: A Systematic Review.* **35**, 356-373 (2019)
- [2] Florence, M. & Bolliger, D. Engagement matters: Student perceptions on the importance of engagement strategies in the online learning environment. *Online Learning.* **22**, 205-222 (2018)
- [3] Shuang, G., Law, K. & Niu, B. Investigating self-directed learning and technology readiness in blending learning environment. *International Journal Of Educational Technology In Higher Education.* **16**, 1-22 (2019)
- [4] Emtinan, A. Predicting student satisfaction and perceived learning within online learning environments. *Distance Education.* **40**, 133-148 (2019)
- [5] Chen Xin, L. Breslow , AND J. Deboer, Analyzing productive learning behaviors for students using immediate corrective feedback in a blended learning environment. *Computers & Education.* **117** pp. 59-74 (2018)
- [6] Priyadarshini, D. & Bandyopadhyay, S. Blended learning to improve quality of primary education among underprivileged school children in India. *Education And Information Technologies.* **24**, 1995-2016 (2019)
- [7] Lin, W., Huang, Y. & Omar, M. Analysis of blended learning model application using text mining method. *International Journal Of Emerging Technologies In Learning (iJET).* **16**, 172-187 (2021)
- [8] Saida, A., Salha, S. & Khlaif, Z. Designing quality e-learning environments for emergency remote teaching in coronavirus crisis. *Interdisciplinary Journal Of Virtual Learning In Medical Sciences.* **11**, 135-137 (2020)
- [9] Juanjuan, C., Wang, M., Kirschner, P. & Tsai, C. The role of collaboration, computer use, learning environments, and supporting strategies in CSDL: A meta-analysis. *Review Of Educational Research.* **88**, 799-843 (2018)
- [10] Almelhi Abdullah, M. Effectiveness of the ADDIE Model within an E-Learning Environment in Developing Creative Writing in EFL Students. *English Language Teaching.* **14**, 20-36 (2021)
- [11] Babatunde, A. & Soykan, E. Covid-19 pandemic and online learning: the challenges and opportunities. *Interactive Learning Environments.* **31**, 7321-7338 (2021)
- [12] Barrot, J., Llenares, I. & Del Rosario, L. Students' online learning challenges during the pandemic and how they cope with them: The case of the Philippines. *Education And Information Technologies.* **26**, 7321-7338 (2021)
- [13] Bibhu Mr & S. K. Yadav, .. High Availability Of Resource In Cloud Computing Technology: Review,Issues And Challenges. *International Journal Of Engineering Trends And Technology.* **68** pp. 1-34 (2020)
- [14] Izotov Oleg Albertovich & A. L. Kuznetsov, .. Development prospects of technologies of grouped cargo containerization. *VESTNIK OF ASTRAKHAN STATE TECHNICAL UNIVERSITY SERIES MARINE ENGINEERING AND TECHNOLOGIES.* **2020** pp. 1-140 (2020)
- [15] Hu Shuyu, .. Research on Data Acquisition Algorithms Based on Image Processing and Artificial Intelligence. *International Journal Of Pattern Recognition And Artificial Intelligence.* **34** pp. 06 (2020)
- [16] Lee Chi Yuan, Shen, C., Lee, S., Chiu, C. & Lin, H. Real-time microscopic monitoring of temperature and strain on the surface of magnesium hydrogen storage tank by high temperature resistant flexible integrated microsensor. *International Journal Of Hydrogen Energy.* **47** pp. 25-12815 (2022)
- [17] Howard, M. Experiential learning—a systematic review and revision of Kolb's model. *Interactive Learning Environments.* **28**, 1064-1077 (2020)
- [18] Sujit, K., Wotto, M. & Belanger, P. E-learning, M-learning and D-learning: Conceptual definition and comparative analysis. *E-learning And Digital Media.* **15**, 191-216 (2018)
- [19] Shivangi, D. Online learning: A panacea in the time of COVID-19 crisis. *Journal Of Educational Technology Systems.* **49**, 5-22 (2020)
- [20] Vandana, S. & Thurman, A. How many ways can we define online learning? A systematic literature review of definitions of online learning (1988-2018). *American Journal Of Distance Education.* **33**, 289-306 (2019)

*Edited by:* Mudasir Mohd

*Special Issue:* Scalable Computing in Online and Blended Learning Environments: Challenges and Solutions

*Received:* May 16, 2023

*Accepted:* Jul 17, 2023







## APPLICATION OF IMPROVED APRIORI ALGORITHM IN INNOVATION AND ENTREPRENEURSHIP ENGINEERING EDUCATION PLATFORM

**Abstract.** The implementation of innovation and entrepreneurship education is inseparable from professional education, so it is important for the rich data in the education platform to mine the connection between professional courses and between grades and courses. The study of association rule algorithm based on education data mining improves the time performance efficiency and accuracy of Apriori algorithm. The study improves the time efficiencies of Apriori algorithm by maintaining Map table and splitting transaction database; the accuracy is improved by using mixed criteria to measure the accuracy and filtering deformation rules based on the inference of confidence. The results of the validation of the time efficiency of the algorithm show that the running time of the improved algorithm in solving frequent itemsets is improved by about 93.86%, 92.48% and 92.76%, respectively, compared with the other three algorithms. The running time of the algorithm for generating frequent itemsets of all orders is about 91.35 ms, which is 66.13% and 83.72% better than the Apriori algorithm and AprioriTid algorithm, respectively. The mining results of student examination data based on the education platform are reasonable and practical, which are of good practical significance for the innovation and entrepreneurship engineering education platform to develop training plans and improve teaching quality. is assumed.

**Key words:** Engineering education platform; innovation and entrepreneurship practice; association rules; data mining; Apriori algorithm

**1. Introduction.** With the advancement of computer technology and database technology, the need for techniques to perform data mining and provide information for decision making based on massive data has continued to grow in various industries [1]. Association rules are important indicators in data mining to reflect the implicit connection between multiple transactions and are being used in the retail industry, educational data mining, medical industry, and financial industry [2, 3]. Association rules are also presented in various forms in the education industry, including the connection between different course grades and the effect of course placement order on overall grades [4]. Among them, analyzing students' test scores and mining the inter-course correlations and learning status information reflected by students' grades can be of great help to students' course selection planning and teachers' teaching according to their needs. There is abundant data in the Innovation and Entrepreneurship Engineering Education Platform, including students' basic information, usual course performance, examination results, stage statistics, etc. These data contain some very valuable information, such as students' mastery of the content they have learned. Therefore, using association rule algorithms to mine the relationship between students' examination results of each course and the relationship between courses is important for activities such as the development of students' course plans and the scheduling of learning progress [5, 6]. However, most domestic universities currently lack a more scientific and in-depth detailed study of the massive data of students' performance. Therefore, the study makes improvements based on the Apriori algorithm and uses the improved algorithm to mine association rules on student data within the database of the education platform, aiming to mine the relationship between students' examination results and the relationship between courses. It helps teachers to develop a curriculum plan that meets students' needs according to their four-year mastery of university courses and the characteristics of engineering education, so as to facilitate the teaching of students' innovation and entrepreneurship practice from shallow to deep and step by step.

**2. Related Work.** With the increasing competition in various industries, data mining techniques are extensively used in diverse fields and have achieved good research success. Yang et al. addressed database technology in data mining, studied the underlying theory and methods of databases, and processed and analyzed data stored based on databases with the aim of improving the usability and popularity of database technology [7]. Amin et al. medical researchers developed prediction models using combination of distinct features and

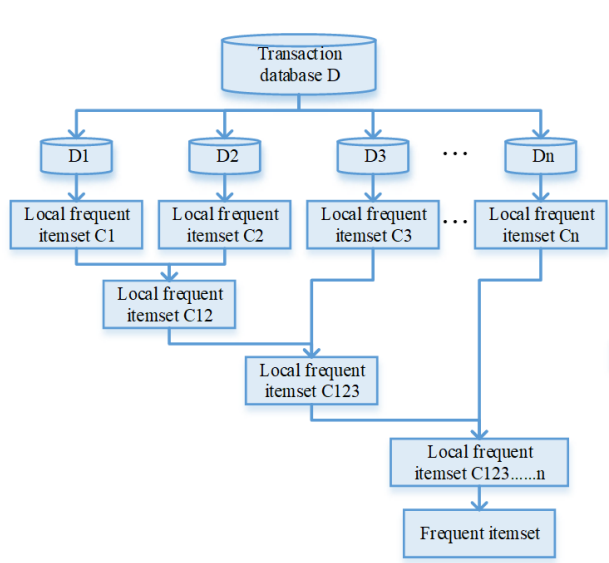
seven classification techniques for identifying salient features and data mining techniques that can enhance the success of predicting cardiovascular diseases and obtained 87.4% accuracy in predicting heart disease [8]. Yun et al. researchers proposed a visualization-aided decision-making system based on data mining techniques for industrial applications. and analyzed the architecture of the system in the context of a practical data mining technique case study verified its effectiveness and robustness [9]. Francis and Babu proposed a new prediction algorithm for predicting student exam results by combining clustering methods and data mining and evaluated and verified the accuracy of this hybrid data mining algorithm in real time using a dataset of students from various academic disciplines in a higher education institution [10]. Ayatollahi et al. compared the performance of artificial neural network and support vector machine-based data mining methods for cardiovascular disease prediction, the latter obtained 16.71 Hosmer-Lemehue test results and up to 92.23% sensitivity, while the area under their subjects' working characteristic curve was larger than that in the neural network model, indicating that support vector machines are more suitable for cardiovascular disease prediction [11].

Finding frequent itemsets is the biggest step in the process of mining association, and one of the most classic algorithms is Apriori algorithm. This algorithm is mainly used for mining single-dimensional binary association rules and has obtained good research results in many fields. Xie X and other scholars proposed a risk prediction and factor risk analysis method fused with Drosophila optimization algorithm and general regression neural network algorithm for coal and gas protrusion accidents, and Apriori algorithm was used to mine the hazard data of accidents. The proposed algorithm was applied to a coal mine to obtain 100% accuracy of accident risk level prediction [12]. Ünvan successively used Apriori algorithm and FP Growth algorithm to analyze a dataset containing 225 products and used FP Growth algorithm to find the top ten rules based on confidence values, according to which supermarkets can make product The supermarkets can adjust the product location based on these rules to increase the product sales and supermarket revenue [13]. Zhan et al. proposed a systematic method for linking customer knowledge and innovative product development in a data-driven settings. The method was used for data mining to obtain customer requirements through Apriori algorithm, correlation rules and decision tree methods. The results proved the validity and usefulness of the proposed method [14]. Ali et al. used the Apriori algorithm in data mining to investigate recovery and mortality factors in a schistosomiasis pretreatment dataset collected from Hubei, China, and evaluated in different tools and models to obtain generative rules with minimum support and minimum confidence indicating higher than 90% and also identified properties indicative of individual recovery and mortality: body mass index, nutrition, degree of ascites, etc., to provide professionals with More accurate guidance [15]. Jing proposed a personalized tourist route intelligent recommendation method based on association rules, which determines the range of tourist attractions by attribute clustering, then extracts the features of attractions, tourists and tourists' interest points by using association rule algorithm to complete the personalized classification of tourist routes, and finally calculates the similarity of tourist routes by dynamic and static attributes and outputs the top attractions with the highest probability. The method obtains 98.5% accuracy, 97% recall and 6s running time in simulation experiments [16].

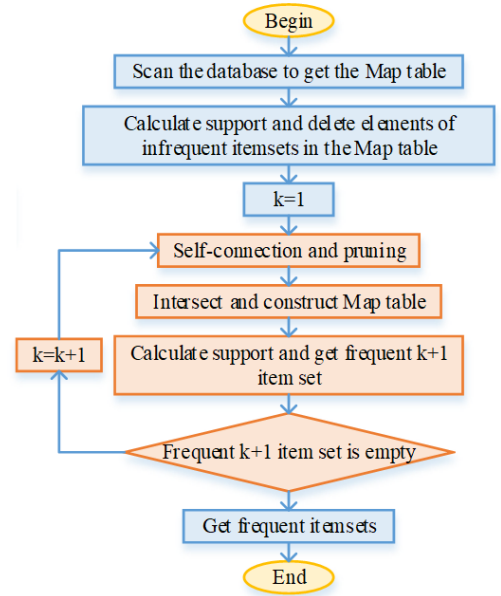
Comprehensive domestic and foreign research scholars through data mining and Apriori algorithm can be found that although Apriori algorithm has good performance in data mining, but there are also certain defects, to apply it to the analysis of student data in the innovation and entrepreneurship engineering education platform, the first need to make improvements for its deficiencies. Therefore, the study improves the time efficiency of Apriori algorithm by maintaining Map table and filtering deformation rules, and applies it to the association rule mining of students' examination results.

### **3. Improved Apriori algorithm for innovation and entrepreneurship engineering education platform.**

**3.1. Time efficiency study of improved Apriori algorithm.** The common association rule algorithms used for educational data mining are Apriori algorithm, Partition algorithm, FP-Growth algorithm, etc [17]. In the generating process of frequent item set by Apriori algorithm, the candidate frequent item set decreases with the increase of order, and the scanning transaction database still needs to traverse all transaction sets and all transactions in the set, which can greatly slow down the algorithm's time efficiency [18]. The improved AprioriTid algorithm reduces the number of scanned transaction sets by constructing a Tid table, but the large quantity of item sets included in the Tid table transactions at the beginning of the algorithm still causes time



(a) Flowchart for reducing the number of transaction database scans



(b) Flowchart for reducing the number of scanned transactions

Fig. 3.1: Time efficiency improvement of Apriori algorithm

wastage. Therefore, the study addresses the shortcomings of Apriori algorithm and AprioriTid algorithm and makes corresponding improvements to them in respect of time performance for the number of transactions in the transactional database and the frequency of transaction database scans. In order to reduce the quantity of transactions scanned in the database, the improved algorithm replaces the operation of scanning the database by maintaining a Map table each time when counting the number of supports for a certain frequent itemset. The number of scans of the database is reduced by splitting the transactional database into several disjoint parts. The exact process of improving time performance is shown Figure 3.1. In Figure 3.1(b), the transactions of each partition are scanned in turn and the local frequent set and its support number are obtained, and then the local frequent set is updated to the candidate frequent set. When all partitioned transactions are scanned, all candidate frequent sets are updated and the final frequent set is determined based on their support numbers and the total number of transactions. To verify the feasibility of the improved algorithm, suppose there is a transaction database with  $m$  records, where the number of transactions contained in the first  $i$  record is  $N_{R[i]}$ . When using the Apriori algorithm for frequent item set solving, the count of transactions  $Num_{[j]}$  to be scanned for the item set of the  $j^{th}$  stratum is calculated as shown in Equation (1).

$$Num_{[j]} = N_{I[j]} \times \sum_{i=0}^{m-1} N_{R[i]} \quad (3.1)$$

In Equation 3.1, assume that the solved frequent itemset has a total of  $k$  order and the sample set of frequent terms contained in the  $j^{th}$  order is  $N_{I[j]}$ . The total quantity of transaction comparisons  $N_{TOTAL}$  is calculated as shown in Equation 3.2.

$$Num_{Total1} = \sum_{j=1}^k Num_{[j]} \quad (3.2)$$

When using the AprioriTid algorithm for support figure calculation of candidate frequent item sets, the corresponding Tid table also has  $k$  order. Assume that there are  $n_{[j]}$  records in the Tid table of order  $j$  and the number of itemsets contained in the  $i$  record is  $N_{TID\_I[i]}$ . Therefore, the number of itemsets in the Tid table scanned by the candidate frequent itemsets at the  $j$  level  $Num_{Tid\_Scan[j]}$  is calculated as shown in Equation 3.3.

$$Num_{Tid\_Scan[j]} = N_{I[j]} \times \sum_{i=0}^{n_{[j]}} N_{Tid_I[i]} \quad (3.3)$$

According to equation 3.3, the corresponding total number of item set comparisons can be obtained from  $N_{TOTAL2}$ , and its computational expression is shown in equation 3.4.

$$Num_{Total2} = \sum_{j=1}^k Num_{Tid\_Scan[j]} \quad (3.4)$$

Instead of scanning the original transaction database, the Map table which corresponds to the candidate itemsets is scanned when the support figure is calculated using the improved Apriori algorithm. The candidate frequent item set number in all strata is shown in Equation 3.5.

$$N_{Total_I} = \sum_{j=1}^k N_{I[j]} \quad (3.5)$$

Since a candidate frequent item set of order  $k$  consists of two frequent item sets of order  $k-1$ , and assuming that the number of IDs contained in the first frequent item set of  $i$  is  $N_{Tid[i]}$  and that candidate frequent item set consists of the first and the first  $Num_{Tid\_Scan[j]}$  item sets, the number of comparisons of the support number of a candidate frequent item set and the total number of comparisons are calculated as shown in Equation 3.6.

$$\begin{cases} N_{Com[i]} = \max(N_{Tid[i1]}, N_{Tid[i2]}) \\ Num_{Total3} = N_{Com[i]} \times N_{Total_I} \end{cases} \quad (3.6)$$

Thus the total number of comparisons of the three algorithms can be approximated by the comparisons of  $\sum_{i=0}^{m-1} N_{R[i]}$ ,  $\sum_{j=0}^{n_{[k]}-1} N_{Tid_I[j]}$  and  $\max(N_{Tid[i1]}, N_{Tid[i2]})$ . In the case of a given transaction database,  $\sum_{i=0}^{m-1} N_{R[i]}$  is a constant value and the AprioriTid algorithm shows that  $n_{[k]}$  is a value less than  $m$  and decreases as the stratum  $k$  increases. Therefore, Apriori algorithm is suitable when there are few records in the database and AprioriTid algorithm is suitable when there are more records. For,  $\max(N_{Tid[i1]}, N_{Tid[i2]})$  also decreases as the hierarchy  $k$  increases, and  $N_{Tid[i]}$  is always smaller than  $m$ , so  $\max(N_{Tid[i1]}, N_{Tid[i2]})i$  is smaller than  $\sum_{i=0}^{m-1} N_{R[i]}$ . When the data is more random, each candidate frequent item set is present in fewer records, i.e.,  $\max(N_{Tid[i1]}, N_{Tid[i2]})$  is smaller. When the data are more similar, it is not possible to visually compare the sizes of  $\sum_{j=0}^{n_{[k]}-1} N_{Tid_I[j]}$  because the size of the two cannot be determined. The comprehensive appeal discussion shows that the improved algorithm is feasible and will have some improvement in time efficiency compared to Apriori algorithm and AprioriTid algorithm.

**3.2. Accuracy study of improved Apriori algorithm.** Commonly used accuracy measures for association rule algorithms include support-confidence, usefulness and validity. Using only two criteria, support and confidence, to measure the results of association mining often generates a large number of useless rules, wrong rules and redundant rules [19]. Considering the advantages and disadvantages of different measures, the study uses a mixture of measures for association rule metrics. The degree of action can be used to measure the degree of influence of the occurrence of the antecedent piece in the association rule on the subsequent piece, and is therefore also known as the degree of relevance, which is calculated as shown in Equation 3.7.

$$\mathbf{Importance}(A \Rightarrow B) = \frac{Support(A \cup B)}{Support(A) * Support(B)} \quad (3.7)$$

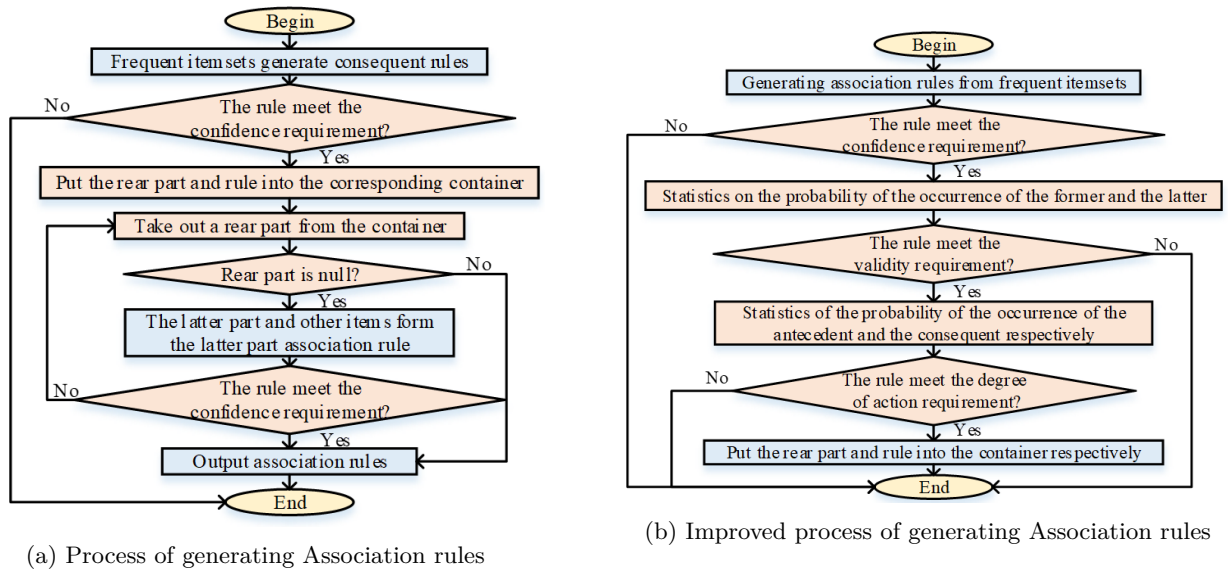


Fig. 3.2: The process of generating affiliation rules before and after improvement

In Equation 3.7, **Importance**( $A \Rightarrow B$ ), represents the relevance of the affiliation rule ( $A \Rightarrow B$ ) and  $Support(A)$  stands for  $A$ 's support level. **Importance**( $A \Rightarrow B$ ) = 1 denotes that  $A$  and  $B$  are mutually independent, and the association rule  $A \Rightarrow B$  is called irrelevant rule. **Importance**( $A \Rightarrow B$ ) > 1 Indicates a positive correlation between  $A$  and  $B$ , i.e., the possibility of  $A$  increases the likelihood of  $B$  occurring, and the association rule  $A \Rightarrow B$  is referred to as a positive correlation rule. However, the correlation still has some limitations, so in order to have a more comprehensive understanding of the impact caused by the occurrence of the item set  $A$  on the occurrence of the item set  $B$ , the concept of effective degree emerged, which is calculated as shown in Equation 3.8.

$$\text{Validity}(A \Rightarrow B) = P(A \cup B) - P(\bar{A} \cup B) \tag{3.8}$$

In Equation 3.8,  $P(A \cup B)$  represents the probability that both the item set  $A$  and the item set  $B$  appear in the transaction database, while  $P(\bar{A} \cup B)$  represents the probability that the item set  $A$  does not appear in the database and the item set  $B$  appears in the database. To improve the accuracy of the association rules, improvements are made to address the shortcomings of the support-confidence measurement process. The process of generating association rules before and after the improvement is shown in Figure 3.2.

As can be seen in Figure 3.2(b), the specific improvement steps are: firstly, we judge the support and confidence level of the generated affiliation rules, and those that meet the confidence requirements enter the next process; then we judge the validity, and delete the association rules that are more likely to appear in the absence of the antecedent; finally, we judge the usefulness, and further delete those association rules whose appearance of the antecedent will lead to a lower probability of the appearance of the consequent. The second improvement is carried out for the deformation rules of association rules. The confidence degree is calculated as shown in Equation 3.9.

$$\text{Confidence}(A \Rightarrow B) = \frac{\text{Support}(A \cup B)}{\text{Support}(A)} \tag{3.9}$$

In Equation 3.9,  $\text{Confidence}(A \Rightarrow B)$  represents the confidence of the association rule ( $A \Rightarrow B$ ). According to Equation 3.7, we can know that assuming the existence of the association rule  $R_5 \Rightarrow R_1 \wedge R_2 \wedge R_3 \wedge R_4$ , its

confidence level is shown in Equation 3.10.

$$\mathbf{Confidence}(R_5 \Rightarrow R_1 \wedge R_2 \wedge R_3 \wedge R_4) = \frac{\mathit{Support}(R_5 \cup R_1 \wedge R_2 \wedge R_3 \wedge R_4)}{\mathit{Support}(R_5)} \quad (3.10)$$

In Equation 3.10,  $\mathit{Confidence}(R_5 \Rightarrow R_1 \wedge R_2 \wedge R_3 \wedge R_4)$  is the confidence level of the association rule  $R_5 \Rightarrow R_1 \wedge R_2 \wedge R_3 \wedge R_4$ . And the confidence of the deformation rule  $R_5 \wedge R_4 \Rightarrow R_1 \wedge R_2 \wedge R_3$  of this rule  $\mathit{Confidence}(R_5 \wedge R_4 \Rightarrow R_1 \wedge R_2 \wedge R_3)$  is shown in Equation 3.11.

$$\mathit{Confidence}(R_5 \wedge R_4 \Rightarrow R_1 \wedge R_2 \wedge R_3) = \frac{\mathit{Support}(R_5 \wedge R_4 \cup R_1 \wedge R_2 \wedge R_3)}{\mathit{Support}(R_5 \wedge R_4)} \quad (3.11)$$

$R_5 \wedge R_3 \wedge R_4 \Rightarrow R_1 \wedge R_2$  The confidence level of  $\mathit{Confidence}(R_5 \wedge R_3 \wedge R_4 \Rightarrow R_1 \wedge R_2)$ , another deformation rule of the association rule  $R_5 \Rightarrow R_1 \wedge R_2 \wedge R_3 \wedge R_4$ , is shown in Equation 3.12.

$$\mathit{Confidence}(R_5 \wedge R_3 \wedge R_4 \Rightarrow R_1 \wedge R_2) \quad (3.12)$$

The confidence level of for , another deformation rule of the association rule , is shown in Equation 3.13.

$$\mathit{Confidence}(R_5 \wedge R_2 \wedge R_3 \wedge R_4 \Rightarrow R_1) = \frac{\mathit{Support}(R_5 \wedge R_2 \wedge R_3 \wedge R_4 \cup R_1)}{\mathit{Support}(R_5 \wedge R_2 \wedge R_3 \wedge R_4)} \quad (3.13)$$

According to the related concept of support, it is known that the association rule has equal support with its three deformation rules, and there is a relationship as shown in Equation 3.14.

$$\begin{aligned} \mathit{Support}(R_5) &\geq \mathit{Support}(R_5 \wedge R_4) \\ &\geq \mathit{Support}(R_5 \wedge R_4 \wedge R_3) \\ &\geq \mathit{Support}(R_5 \wedge R_4 \wedge R_3 \wedge R_2) \end{aligned} \quad (3.14)$$

The confidence relationship of association rules and their deformation rules can be obtained by combining Equation 3.10 to Equation 3.14, as shown in Equation 3.15.

$$\begin{aligned} \mathit{Confidence}(R_5 \Rightarrow R_1 \wedge R_2 \wedge R_3 \wedge R_4)_{\min} \\ &\leq \mathit{Confidence}(R_5 \wedge R_4 \Rightarrow R_1 \wedge R_2 \wedge R_3) \\ &\leq \mathit{Confidence}(R_5 \wedge R_4 \wedge R_3 \Rightarrow R_1 \wedge R_2) \\ &\leq \mathit{Confidence}(R_5 \wedge R_4 \wedge R_3 \wedge R_2 \Rightarrow R_1) \end{aligned} \quad (3.15)$$

In Equation 3.15,  $\mathit{Con}_{\min}$  is the minimum confidence level. Equation 3.15 shows that the morphing rules of an association rule are all strong correlation rules that satisfy the confidence measure[20]. When a powerful association rule is obtained, its similar deformation rules must be powerful association rules, but this part of rules is not our target rules, so we need to filter out this part of deformation rules in the process of producing association rules.

**3.3. An Improved Apriori algorithm based on the education platform student performance analysis method.** In the innovation and entrepreneurship engineering education platform, Students' performance in examinations is not only an overwhelming measure of teaching quality, but also an important basis for guiding students to properly select and complete the corresponding innovation and entrepreneurship courses [21]. The structure of student performance analysis application system is shown in Figure 3.3.

In Figure 3.3, the study divides the student performance analysis application into three modules according to different user needs: student performance analysis module, individual student performance analysis module, and teacher course analysis module. The student grade analysis module is mainly used to mine the correlations among students' course grades in the four years of college and store the obtained rules in the rule base. The process of generating association rules according to the improved Apriori algorithm is as shown in Figure 3.4.

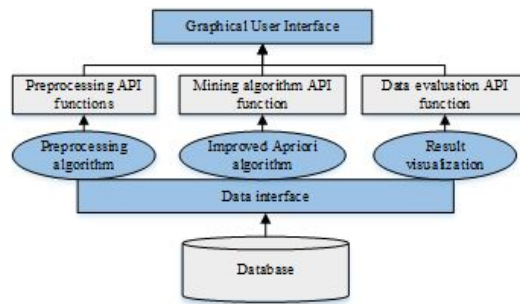


Fig. 3.3: Structure chart of student performance analysis application

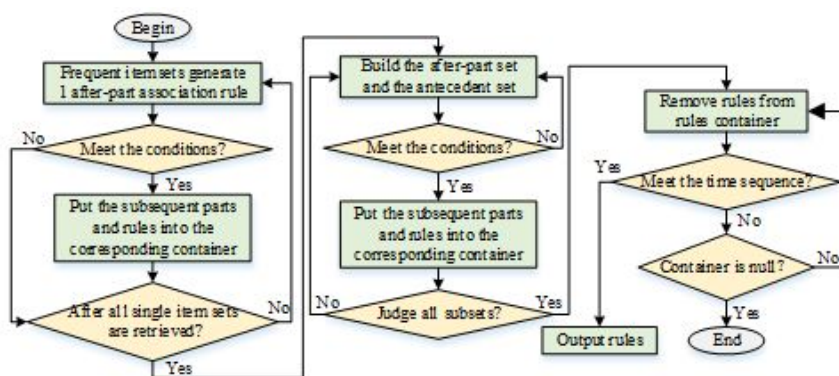


Fig. 3.4: The process of generating association rules according to the improved Apriori algorithm

As shown in Figure 3.4, a single item is selected as a posterior in the input frequent item set at a time, and it is judged whether the association rule satisfies the relevant conditions, and then the posterior items that satisfy the conditions are put into the posterior container, and the formed association rule is put into the corresponding rule container. After that, the items in the post-item container are taken in turn to form the post-item set with other items in the frequent item set, and the other items in the item set form the pre-item set. Finally, take the elements from the rule container in turn and judge whether they satisfy the temporal order, and print the rules that satisfy the condition, otherwise discard the rule. When the container of association rules is empty it means the generation of association rules is completed. The student performance analysis module is an analysis of all students' performance. Students can view the relationship between the grades of each course, but it is difficult to visually identify the rules related to them from the results. Therefore, the study uses the course exam results entered by students as the antecedent of the correlation rules in the individual student performance analysis module, and predicts students' performance in subsequent courses based on the generated rules. Based on the prediction results, students can have a clearer idea of which courses they need to focus on in their subsequent courses, so that they can make scientific and effective study plans in advance. In university teaching, the study of public basic courses is often carried out through large classes, while some subjects with stronger specialization, such as those in the innovation and entrepreneurship engineering education platform, are mostly taught in small classes. The teacher curriculum analysis module can analyze the teacher curriculum according to the connection between students' performance in each subject and carry out scientific class teaching, which is more conducive to teachers' teaching according to their ability.

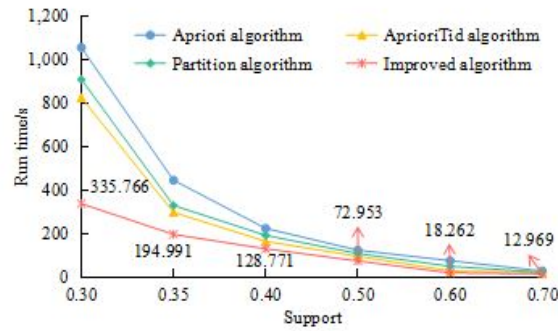


Fig. 4.1: The time of solving frequent itemsets by different algorithms with small transaction differences

#### 4. Results of the improved Apriori algorithm in the innovation and entrepreneurship engineering education platform.

**4.1. Performance analysis of Apriori algorithm before and after optimization.** The performance comparison experiments before and after the optimization of Apriori algorithm are based on the verification of time efficiency improvement, firstly, by conducting time complexity comparison experiments on Apriori algorithm, AprioriTid algorithm, Partition algorithm and the improved algorithm. The time complexity comparison experiments are conducted in three times, including time-efficient effectiveness in the situation of small differences in transactions, time-efficient effectiveness in the situation of large differences in transactions, and time-efficient effectiveness in the generation of frequent itemsets of each order. The data in Experiment 1 are 1000 data excerpted from the UCI dataset and the mushroom.dat data in PUMSB, where the differences between each transaction are small. By varying support settings, the time required for each algorithm to solve the set of frequent items under different thresholds was counted, and the experimental results obtained are shown in 4.1.

In Figure 4.1, both the AprioriTid algorithm and the Partition algorithm, which are derived from the improved Apriori algorithm, have a certain degree of reduction in the time needed to solve the frequent itemset compared to the Apriori algorithm, and the algorithms are more time efficient. The improved algorithm has less running time compared to other algorithms, especially in the range of support threshold of 0.30 0.40, and this time efficiency improvement is more obvious. Experiment 2 selected 1000 data from the T10I4D100K dataset to verify the time-efficient of the improved algorithm in the presence of large transaction differences. Due to the significant differences between transactions in the experimental data, a smaller minimum support needs to be set in the setup, otherwise the frequent item set cannot be obtained. The time required by the four algorithms to solve the frequent itemset under different thresholds is shown in Figure 4.2.

As shown in Figure 4.2, the improved algorithm improves the running time by about 93.86%, 92.48% and 92.76% compared to the Apriori algorithm, AprioriTid algorithm and Partition algorithm when the support threshold is 0.010. The improved algorithm improved the average running time by about 91.93%, 81.43% and 88.88% compared to the Apriori algorithm, AprioriTid algorithm and Partition algorithm for different support thresholds, respectively. The runtime ratio of the improved algorithm to the other three algorithms in Experiment 2 is larger compared to Experiment 1. The reason for this analysis is that with significant differences between transactions, the number of transaction IDs contained in the Map table corresponding to each item set is relatively small, so the comparison times are decreased and the effectiveness of solving frequent item sets is improved. Experiment 3 is to count the running time of different algorithms during the generation of frequent itemsets of each order under the support threshold of 0.3 and 0.5, and the results are presented in Figure 4.3.

In Figure 4.3(a) and Figure 4.3(b), the difference between the running time of the Apriori algorithm for generating frequent itemsets of each order is not significant, while the runtime of the AprioriTid algorithm for



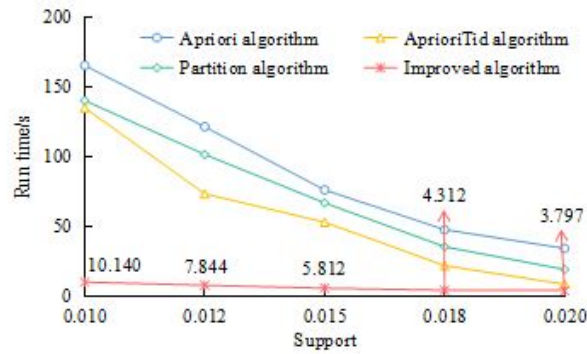
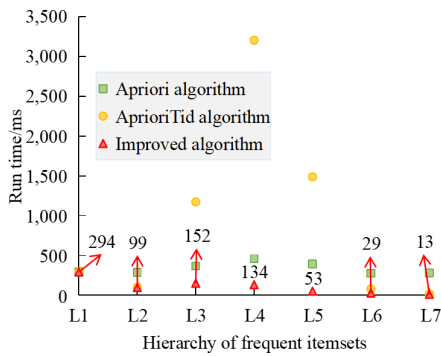
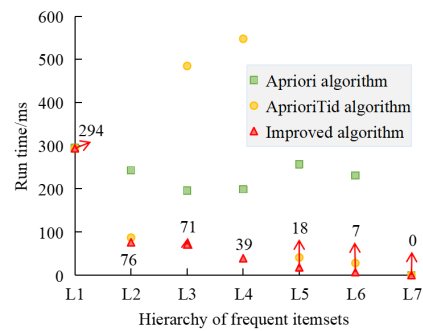


Fig. 4.2: The time of solving frequent itemsets with different algorithms in the case of large transaction differences



(a) Time to generate frequent itemsets of each order when the support is 0.3



(b) Time to generate frequent itemsets of each order when the support is 0.5

Fig. 4.3: Time to generate frequent itemsets of each order when support is 0.3 and 0.5

generating L3 itemsets and L4 itemsets is much larger than that of the Apriori algorithm. However, as the class of frequent itemsets increases, the transactions in the AprioriTid algorithm gradually decrease and the algorithm running time is significantly lower than that of the Apriori algorithm. Regardless of the support threshold of 0.3 or 0.5, the running time efficiency of the improved algorithm is always significantly higher than that of the Apriori algorithm and the AprioriTid algorithm throughout the generation of frequent itemsets. The average running times of the improved algorithm, Apriori algorithm and AprioriTid algorithm for each order of frequent item set generation were about 91.35 ms, 269.72 ms and 561.00 ms.

**4.2. Effectiveness of Apriori algorithm for mining student achievement data in educational platform.** The study selected all student grades of three graduated grades of software engineering majors in the database of the innovation and entrepreneurship engineering education platform as sample data, and conducted data preprocessing operations such as checking, filling, cleaning, and conversion. In the data preprocessing, considering the excessive variability of each student’s choice of elective courses, only the students’ compulsory course grades were selected and only the grades of students’ first exams were used. Considering that string matching and storing the Chinese names of courses would cause a large memory overhead, numeric codes are used instead of the Chinese names of courses. The probability of having duplicate data in the resulting set of



(a) The number of rules generated in two cases when the support is 0.2



(b) The number of rules generated in two cases when the support is 0.3

Fig. 4.4: Number of rules generated by time decision and no time decision under different support

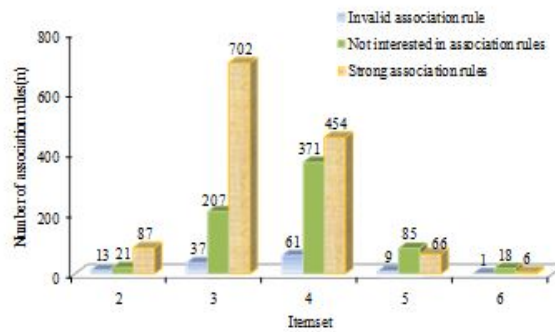


Fig. 4.5: The quantity of frequent itemsets of each order generated by the improved algorithm for data mining

frequent items is too low due to too large a span of student grades, so five grades of A, B, C, D, and E are used instead of percent test scores, and the weighted average score is denoted by the letter Z. When a student misses an exam in a course, if the student has a make-up exam grade, the make-up exam grade is used as the exam grade, otherwise it is automatically classified as an E grade. To analyze the effect of the academic year taken in the course on the mining results, the study collected the result data obtained without considering the academic year taken in the course and considering the academic year taken in the course, as shown in Figure 4.4.

As shown in Figure 4.4, with support degrees of 0.2 and 0.3, the number of association rules generated when considering the academic year taken in the course is much less than that generated when not considering the academic year taken in the course, and the ratio of the number of association rules is about 0.63. Therefore, the study adds the judgment of time to the algorithm for grade analysis, so that some redundant association rules can be removed. The minimum degree of support is set to 0.198 and the minimum confidence level is set to 0.657. The number of frequent item sets of each order generated by the improved algorithm for data mining is shown in Figure 4.5.

As shown in Figure 4.5, invalid association rules and uninteresting association rules account for a large proportion of the results. Some of the invalid association rules can be removed by the validity and usefulness,

Table 4.1: Some Intense Association Rules in Mining Results

Rule Antecedent	Rule Aftereffect	Support of Front Parts	Confidence
2C	5C	0.3333	0.9259
6A	4A	0.2469	0.8250
1A	11A	0.4383	0.8169
3A	ZA	0.3333	0.7593
5C	7C	0.2346	0.7556
10B	1A	0.5370	0.7287
9A	ZA	0.4198	0.7260
4A	1A	0.3148	0.7255
8C	9C	0.4136	0.7075
8A	ZA	0.4198	0.7059

and the uninteresting rules can be removed by the inference formula and course time judgment. Removing these wrong rules can better select the rules that are more beneficial to users for popularization. Some of the strong association rules in the mining results are presented in Table 4.1.

A higher confidence level for the association rule in Table 1 represents a greater degree of association between the preceding and following pieces of the rule, that is, a higher degree of importance of the rule. The study divided the strong association rules into the rules of course scheduling rationality and the rules of subject's influence on overall performance. According to the rules  $2C \Rightarrow 5C$  and  $5C \Rightarrow 7C$  in Table 1, it can be found that if a software engineering student gets a poor grade in course 2, there is a high probability that it will lead to a poor grade in course 5, which will eventually affect the grade in course 7. Therefore, when arranging courses in the education platform, we need to arrange course 2 in the front and focus on improving teaching level to prepare the basis for students to take course 5 and course 7 next. According to the three rules in Table 1  $3A \Rightarrow ZA$ ,  $8A \Rightarrow ZA$  and  $9A \Rightarrow ZA$ , it can be seen that Course 3, Course 8 and Course 9 have a greater impact on students' overall performance, i.e. students with good performance in these subjects have mostly higher overall performance. Therefore, the credit weights of these courses can be appropriately adjusted in the innovation and entrepreneurship engineering education platform as a way to balance the professional level of students.

**5. Conclusion.** The cultivation of innovation and entrepreneurship is an important part of the reform of student cultivation mode in China's universities, and there are high requirements for students' capability and innovation skills in engineering education. Therefore, it is necessary to arrange courses and innovation and entrepreneurship practice activities in a targeted way based on students' ability to master the subject theory. The Apriori algorithm of association rule mining is studied to mine the implicit information in students' performance. Improvements to the Apriori algorithm are made in two aspects: time efficiency and accuracy, specifically by reducing the number of scanned databases and transactions to improve the operational efficiency of the algorithm, and by filtering deformation rules to improve the accuracy. In the validation experiments on the effectiveness of the improved performance, the time required for the improved algorithm to resolve the frequent itemsets is improved by about 93.03% compared with the other three algorithms. The operation time of the improved algorithm to generate frequent itemsets of each order is about 91.35 ms, which is 66.13% and 83.72% better than the Apriori algorithm and AprioriTid algorithm, respectively. However, the algorithm proposed in the study only refers to the existing support and confidence levels in relevant research when generating association rules, and does not delve into the most suitable selection method for support and confidence levels. In the future, further optimization of the algorithm should be carried out for the selection of minimum support and minimum confidence levels, and further research is needed on how to select the best rule among the generated association rules for popularization.

**Fundings.** The research is supported by: 2023 Hunan Provincial Social Science Achievement Evaluation Committee member general funding project “Research on Innovation of Professional Innovation Integration Path under Employment Priority Strategy” (No.: XSP2023JYZ022).

## REFERENCES

- [1] Sunhare, P., Chowdhary, R. & Chattopadhyay, M. Internet of things and data mining: an application-oriented survey. *Journal of King Saud University- Computer and Information Sciences.* (2022)
- [2] Menon, S., Ghoshal, A. & Sarkar, S. Modifying transactional databases to hide sensitive association rules. *Information Systems Research.* **33**, 152-178 (2022)
- [3] Thurachon, W. & Kreesuradej, W. Incremental association rule mining with a fast incremental updating frequent pattern growth algorithm. *IEEE Access.* **9** pp. 55726-55741 (2021)
- [4] Aldowah, H., Al-Samarraie, H. & Fauzy, W. Educational data mining and learning analytics for 21st century higher education: a review and synthesis. *Telematics And Informatics.* **37** pp. 13-49 (2019)
- [5] Ghafar, A. Convergence between 21st century skills and entrepreneurship education in higher education institutes. *Education.* **9**, 218-229 (2020)
- [6] Duan, N., Duan, L. & Lu, H. The current situation of innovation and entrepreneurship education and its optimization countermeasures: A case study of International Journal of Social Science and Education Research. (2021)
- [7] Yang, J., Li, Y., Liu, Q., Li, L., Feng, A., Wang, T., Zheng, X. & Lyu, J. Brief introduction of medical database and data mining technology in big data era. *Journal of Evidence-Based Medicine.* (2020)
- [8] Amin, M., Chiam, Y. & Varathan, K. Identification of significant features and data mining techniques in predicting heart disease. *Telematics and Informatics.* (2019)
- [9] Yun, Y., Ma, D. & Yang, M. Human-computer interaction-based decision support system with applications in data mining. *Future Generation Computer Systems.* (2021)
- [10] Francis, B. & Babu, S. Predicting academic performance of students using a hybrid data mining approach. *Journal of Medical Systems.* (2019)
- [11] Ayatollahi, H., Gholamhosseini, L. & Salehi, M. Predicting coronary artery disease: a comparison between two data mining algorithms. *BMC Public Health.* (2019)
- [12] Xie, X., Fu, G., Xue, Y., Zhao, Z., Chen, P., Lu, B. & And, J. and factors risk analysis based on IFOA-GRNN and apriori algorithms: application of artificial intelligence in accident prevention. *Process Safety And Environmental Protection.* **122** pp. 169-184 (2019)
- [13] Ünvan, Y. Market basket analysis with association rules. *Communications In Statistics-Theory And Methods.* **50**, 1615-1628 (2021)
- [14] Zhan, Y., Tan, K. & Huo, B. Bridging customer knowledge to innovative product development: a data mining approach. *International Journal Of Production Research.* **57**, 6335-6350 (2019)
- [15] Ali, Y., Farooq, A., Alam, T., Farooq, M., Awan, M. & Baig, T. Detection of schistosomiasis factors using association rule mining. *IEEE Access.* **7**, 86108-18611 (2019)
- [16] Jing, Y. An intelligent recommendation method of personalised tour route based on association rules. *International Journal of Reasoning-based Intelligent Systems.* (2023)
- [17] Cantabella, M. Martínez-España R, Ayuso B, Yáñez J A. *Mu Noz A. Analysis Of Student Behavior In Learning Management Systems Through A Big Data Framework.* **90** pp. 262-272 (2019)
- [18] Zhu, S. Research on data mining of education technical ability training for physical education students based on Apriori algorithm. *Cluster Computing.* **22** pp. 14811-14818 (2019)
- [19] Wang, Z., Tian, Q. & Duan, X. Research on the evaluation index system of college students' class teaching quality based on association algorithm. *Cluster Computing.* **22** pp. 13797-13803 (2019)
- [20] Wang, H. & Gao, Y. Research on parallelization of Apriori algorithm in association rule mining. *Procedia Computer Science.* **183** pp. 641-647 (2021)
- [21] Khan, A. & Ghosh, S. Student performance analysis and prediction in classroom learning: a review of educational data mining studies. *Education and Information Technologies.* (2021)

*Edited by:* Mudasir Mohd

*Special Issue on:* Scalable Computing in Online and Blended Learning Environments: Challenges and Solutions

*Received:* May 18, 2023

*Accepted:* Jul 18, 2023



## SPACECRAFT TEST DATA INTEGRATION MANAGEMENT TECHNOLOGY BASED ON BIG DATA PLATFORM

**Abstract.** In this paper, a general test platform for spacecraft data management is designed and constructed. This paper introduces a portable software development environment based on LUA. The technology of space environment data management, comprehensive analysis, parameter correction and visual display of spacecraft is realized. The relationship between continuity, mixed dispersion, variation and indication of remote sensing data is studied. This project uses the integrated Long Short Term Memory network (LSTM) technology to detect anomalies in satellite remote sensing observation data. Give full play to the advantages of laser scanning tunneling microscope in the nonlinear field. The combination of this method and the matrix method can improve the adaptive ability of spacecraft in an operation state to better identify abnormal information in remote sensing data. Experiments show that the algorithm can significantly improve the anomaly detection rate of the system. The system can monitor the front test device and record the data. The method can be connected with the space vehicle's central control and automatic test system. The comprehensive management of the integrated test system of space vehicles is realized.

**Key words:** Spacecraft; Front-end test equipment; Telemetry; Anomaly detection; LSTM

**1. Introduction.** The spacecraft data processing subsystem (DMS), as the "brain" of spacecraft, plays an essential role in space information processing. DMS undertakes remote sensing measurement in aviation systems, such as remote sensing data collection and transmission, remote control command receiving and execution, flight plan maintenance and space flight routine work. Each subsystem on the spacecraft is subject to distributed control management, so the stability and reliability of distributed control management are crucial to the successful operation of the spacecraft. DMS functional testing is significant. Spacecraft face a highly complex space environment when they fly in space [1]. It involves all kinds of particle radiation, electromagnetic radiation, atomic oxygen, temperature changes, etc. Real-time monitoring of the space environment and obtaining relevant space environment information have practical application value and prospects. It provides a good condition and basis for the country's high-life and high-quality artificial satellites. At the same time, the spacecraft data processing subsystem will also provide necessary data support for space weather prediction and scientific research [2]. Promoting the international exchange of space environment detection technology and data in China is of great significance.

Scientific space environment monitoring is significant to studying astrophysics and planetary evolution. There has been a long history in the field of automatic detection of space vehicles. The software and hardware products on the space vehicle have formed a relatively complete automatic detection method library [3]. The system can realize the function check and reliability test of hardware equipment and software. Some scholars have proposed a ground test device for space vehicles based on artificial intelligence. This system has an automatic discriminating ability similar to conditional criteria. The system is applied to the automatic verification test of the power supply and distribution equipment of space vehicles, which can improve the verification test effect of the equipment. In recent years, the automatic detection technology of spacecraft has been dramatically developed in China. The method has been successfully used in space fields such as human-crewed spaceflight and the Beidou satellite. However, the current automated test system design at home and abroad mainly focuses on the execution of spacecraft test sequences. Most related systems are configured at the back-end level of the EGSE system. Little attention has been paid to the automatic operation of the front-end test unit. Although some automatic detection systems have detection devices, they can only complete the command control of the detection devices. It cannot obtain the working condition of the device under test in time. The device has a good coupling relationship with the tested space vehicle. In addition, the system is highly dependent on hardware. The development of a new standardized intelligent test platform can fully meet the testing needs of space agencies. The system platform designed in this paper abstracts the test business of the spacecraft

data management subsystem, thus completing the standardization of the organization structure, test flow and communication protocol of the spacecraft data management subsystem test system [4]. At the same time, introducing the LUA scripting language makes the test program customization more intelligent and standardized. The test bed is also a reconfigurable system. The system can be tested under the imperfect spacecraft data processing subsystem architecture.

## 2. Design of automatic management system.

**2.1. System Software Architecture.** Based on the mature integrated automatic detection technology, an automatic detection technology for spacecraft front-end detection devices is proposed. This system uses Visual Studio software. The software is designed according to the modularization idea [5]. Divide the system functions to be completed into several functional modules. The data interface between each module is straightforward. It has good expansibility in device type and function control. Figure 2.1 shows the architecture of the entire system. It is mainly divided into six parts: 1) The network communication part completes the network interface function of the system. A network connection is provided between the managed front-end test device and the master control master test processor (MTP). Receive the health information sent by the managed device and process the content of the information according to the agreement. Send the control command to the device and get a reply. 2) The data analysis function module retrieves packets sequentially from a valid packet cache queue. According to the designed packet format, the spacecraft identifier, the local time of the sending time of the sender, the device identifier of the sender, the device identifier of the receiver, the information identifier, the cumulative count of the packet sending of the sender, and the data content are analyzed. The identification of this information determines the status of this level. And write the received data into a log file. 3) The instrument control module is dynamically maintained as a data link table. Managed devices can be added or removed via the system's software interface or profile. Device information includes device name, device number, control instruction name, control instruction number, IP address, port number, etc. The network communication module receives the control indicator and analyzes the indicated content. The device generates an indication response according to the communication protocol and returns a message to the network communication module. 4) The program's sequence implementation part reads the instruction sequence selected by the human-machine interface. According to the command number and the device ID, it is sent to the device management module as the corresponding command. You can set the time interval for instruction execution, the number of loops, and boundary conditions. 5) The data file module records the diary data generated by the system in its operation and the data received by the managed device. At the same time, it makes an instant recording. 6) Display/operation module displays the working status of the system. Various types of status information transmitted by the managed device respond to the tester's input. It is a kind of management system software with a man-machine dialogue function.

## 2.2. System module design.

**2.2.1. Product structure module.** According to the four levels of whole, cabin, cabin board and equipment, the prototype is constructed to realize the whole structure of the product. The process prototype is summarized from the system's overall structure, final assembly design pattern, subsystem design pattern and related design documents. The system was produced during the integration process. The creation process is shown in Figure 2.2.

**2.2.2. Spacecraft data transmission module.** The same information is often recorded in multiple places during the AIT process. Such duplication of data leads to "data contradictions." This situation has an impact on the tracking of the process. Therefore, the system should follow the principle of consistency of data sources. Ensure all data is transferred from existing systems or entered via files etc. According to different information systems, the corresponding communication mode and interface are designed in this paper [6]. The system is integrated through network service, ETL, intermediate table synchronization and other methods to determine whether the transmitted data is complete. The data management system itself only transfers data from other data systems. The data outside the work network is entered into the corresponding information system through the physical media. Then use the data interface of the office network to transmit this information. See inter-system data integration relationship Figure 2.3 (image cited by the National Aeronautics

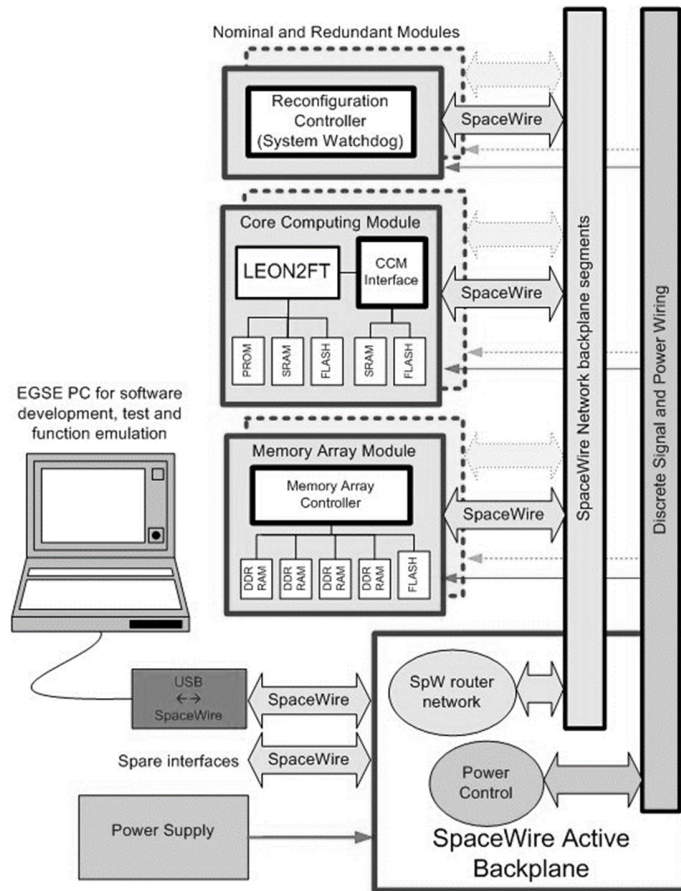


Fig. 2.1: Software module and interface of the front-end detection device AIS.

and Space Administration).

**2.2.3. Space data reconstruction and contact module.** The data reconstruction and association module reconstructs process-based data. Connect it to the "Product Architecture Framework" to complete the data content filling. Figure 2.4 shows the process flow.

The information in the message display system comes from different systems. Benchmarking and versioning the data in the process is necessary to ensure the integrity and accuracy of AIT process data [7]. Keep the same version of data for each dimension. This can effectively prevent the query error caused by the difference in data.

**2.2.4. Data Application Module.** The data module is used for statistics and a summary of the forms of each module. It includes all biased data, custom queries and data mining. The retrievability and scalability of AIT data are realized through the configuration and management of big data such as Impala. Thus, the vertical comparison between the same model and the horizontal comparison between the different models is achieved. For example, it can quickly compare the execution records of each disassembly and site photos in multiple disassembly operations of an important part. The assembly situation and data of the same type of equipment on the same batch of satellites are easily compared.

**3. Anomaly detection of spacecraft remote sensing data based on LSTM.** LSTM model neglects the modal characteristics of remote control commands when establishing remote control commands. This will affect the accuracy of the forecast. This dramatically impacts some problems existing in satellite remote sensing

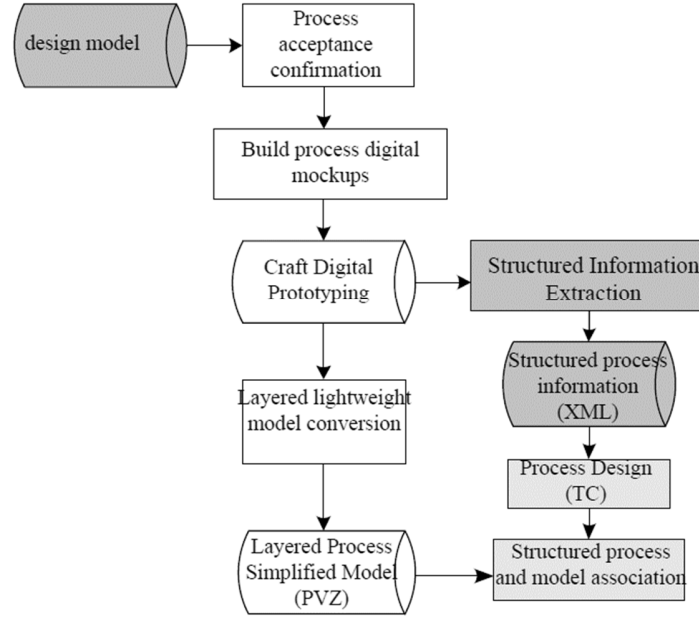


Fig. 2.2: Data generation process based on product structure.

data. This project plans to conduct research from three aspects: 1) Feature extraction of control commands. 2) Multi-LSTM model training. 3) Anomaly detection method based on remote sensing data. This paper presents an algorithm for anomaly detection of remote sensing data using an integrated LSTM prediction model (Figure.3.1).

**3.1. Excavation of Control command pattern.** The algorithm is performed before the LSTM prediction model is started. The first is the preprocessing of these training samples. The method of cluster analysis is used to classify them [8]. Among them, the preprocessing of training data and the mining of remote indication control mode are the focus of this paper.

**3.1.1. Training data preprocessing.** The training set  $R$  can be regarded as a set of  $n$  dimensional vectors at time  $t$ .  $r^t = \{r_1^t, r_2^t, \dots, r_n^t\}$ . Where  $\{r_1^t, r_2^t, \dots, r_{n-1}^t\}$  is the multi-dimensional remote control command at time  $t$ .  $\{r_n^t\}$  is the telemetry data at time  $t$ . Then the expression  $R$  for the training set of telemetry data containing a subsystem is:

$$R = \left\{ \left[ \begin{array}{c} r_1^1 \\ \vdots \\ r_{n-1}^1 \\ r_n^1 \end{array} \right], \left[ \begin{array}{c} r_1^2 \\ \vdots \\ r_{n-1}^2 \\ r_n^2 \end{array} \right], \dots, \left[ \begin{array}{c} r_1^t \\ \vdots \\ r_{n-1}^t \\ r_n^t \end{array} \right], \dots \right\} \quad (3.1)$$

Then  $R$  is reconstructed. Because the telemetry data at the next moment is related to the remote command at the next moment, and the telemetry data at the previous time is related to the corresponding remote command [9]. Therefore, when constructing the input,  $\tilde{R}$  is divided into multiple submatrices  $R_m^*$  containing continuous-time vectors. Each submatrix contains telemetry data and global remote control instructions within  $s_t$  time. Input  $\tilde{R}$  for reconstruction is represented as follows

$$\tilde{R} = \{\{R_1^*\}, \{R_2^*\}, \dots, \{R_m^*\}, \dots\} \quad (3.2)$$



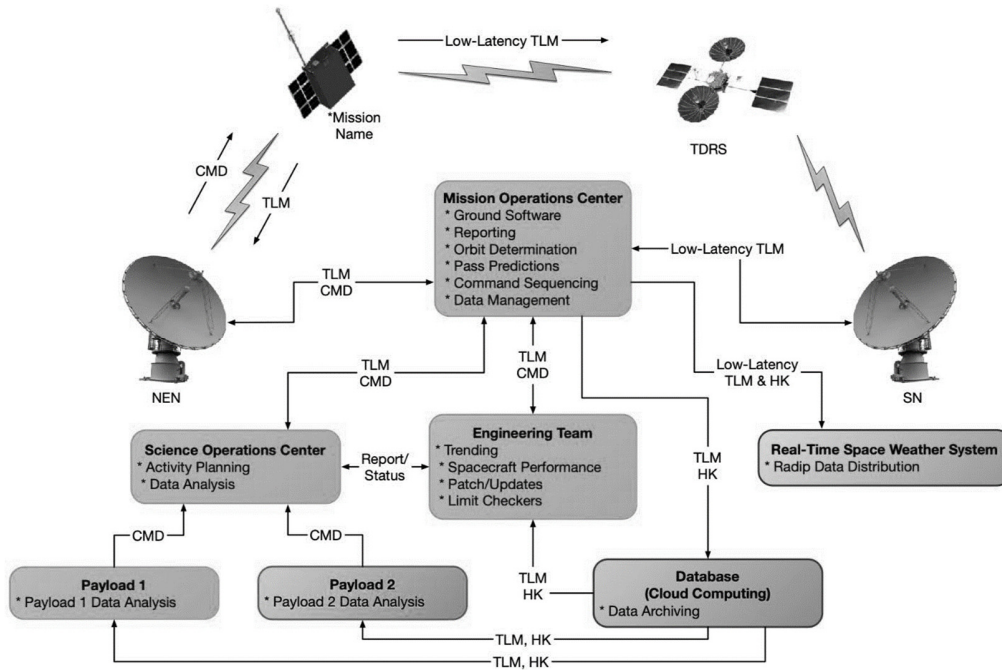


Fig. 2.3: Data integration relationships between different systems.

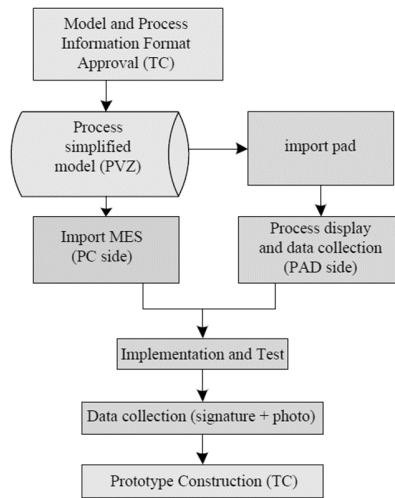


Fig. 2.4: Data reconstruction and association process.

Among them:

$$R_m^* = \left\{ \begin{bmatrix} r_1^m \\ r_2^m \\ \vdots \\ r_n^m \end{bmatrix}, \begin{bmatrix} r_1^{m+2} \\ r_2^{m+2} \\ \vdots \\ r_n^{m+2} \end{bmatrix}, \dots, \begin{bmatrix} r_1^{s_t+m} \\ r_2^{s_t+m} \\ \vdots \\ r_n^{s_t+m} \end{bmatrix} \right\} \quad (3.3)$$

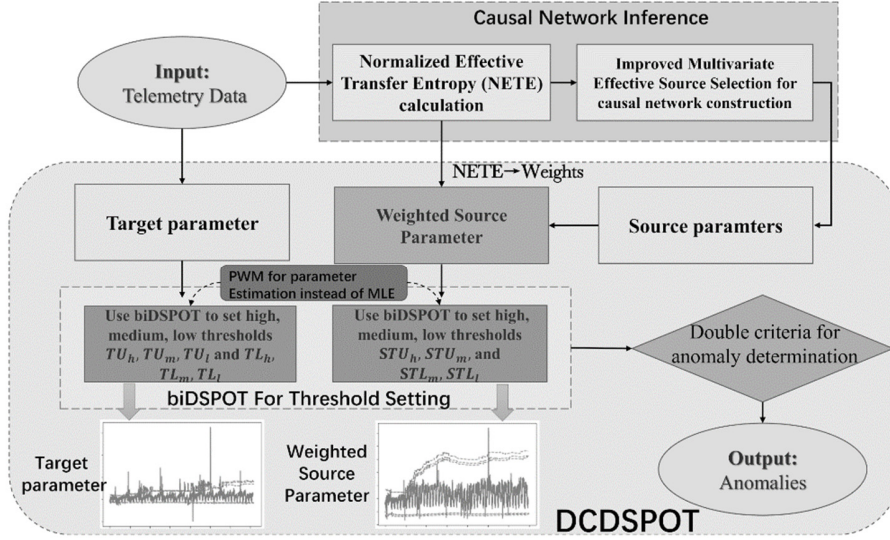


Fig. 3.1: Anomaly detection method of remote sensing data integrated with LSTM prediction model.

**3.1.2. Remote command control mode mining.** Take input  $\tilde{R}$  as an example to extract the remote control instruction  $D_m^*$  in each submatrix  $R_m^*$ . In this way, the control instruction matrix  $D^*$  corresponding to  $\tilde{R}$  is obtained.

$$D_m^* = \left\{ \left[ \begin{array}{c} r_1^m \\ r_2^m \\ \vdots \\ r_{n-1}^m \end{array} \right], \left[ \begin{array}{c} r_1^{m+1} \\ r_2^{m+1} \\ \vdots \\ r_{n-1}^{m+1} \end{array} \right], \dots, \left[ \begin{array}{c} r_1^{s_t+m} \\ r_2^{s_t+m} \\ \vdots \\ r_{n-1}^{s_t+m} \end{array} \right] \right\} \quad (3.4)$$

By calculating the  $S_2$  norm of  $D^*$  submatrix. In this way, a set of control mode feature vector  $\|D^*\|_2$  is obtained.

$$\|D^*\|_2 = \{\|D_1^*\|_2, \|D_2^*\|_2, \dots, \|D_m^*\|_2, \dots\} \quad (3.5)$$

$$\|D_m^*\|_2 = \sqrt{\eta_{\max}(D_m^{*T} D_m^*)} \quad (3.6)$$

Reorder the values in the control mode feature vector  $\|D^*\|_2$ . The median in  $\|D^*\|_2$  is taken as the threshold by dividing it into two sub-training sets. The similar sequence of centralized remote instructions divides the training set into 2 classes.

**3.2. Multi-LSTM prediction model training.** LSTM prediction model respectively trained the two sub-training sets obtained by pattern mining and clustering. For example, the reconstructed training matrix  $R$  is used as input for telemetry data prediction. Let  $t = s_t$ , the prediction process is as follows:

$$R_m^* = \left\{ \left[ \begin{array}{c} r_1^m \\ r_2^m \\ \vdots \\ r_n^m \end{array} \right], \dots, \left[ \begin{array}{c} r_1^{m+t} \\ r_2^{m+t} \\ \vdots \\ r_n^{m+t} \end{array} \right] \right\} \rightarrow \left[ \begin{array}{c} r_1^{m+t+1} \\ r_2^{m+t+1} \\ \vdots \\ \hat{f}(x)^{m+t+1} \end{array} \right] \quad (3.7)$$

$\hat{f}(x)^{m+t+1}$  is the predicted value at time  $m+t+1$ . When constructing the prediction model, the absolute error function (MAE) is the loss function of the LSTM prediction model:

$$MAE(f(x), \hat{f}(x)) = \frac{\sum_{i=1}^m |f(x)_i, \hat{f}(x)_i|}{m} \quad (3.8)$$

$m$  is the number of total predicted values.  $f(x)_i$  is the  $i$  true value.

The optimal solution of LSTM prediction model is Adam. The sensitivity of the method to non-normal changes is improved [10]. To avoid overfitting, the memory modules of each hidden layer cannot enter the LSTM learning. This approach is also known as dropout. LSTM prediction model P and LSTM prediction model Q are obtained after all the sub-training sets are trained.

**3.3. Telemetry data anomaly detection model.** After the training of the LSTM prediction model  $P$  and LSTM prediction model  $Q$ . The test data will be separately passed through each prediction model. Two prediction sequences were obtained [11]. The final prediction sequence is obtained by integrating two groups of prediction sequences according to specific weights.

The test data is reassembled using training set  $\tilde{R}$  as an example [12]. The reconstructed test set was simultaneously predicted by LSTM prediction model  $P$  and LSTM prediction model  $Q$  to obtain two sets of prediction sequences. Telemetry data prediction set  $P, \hat{f}(x)_P = \{\hat{f}(x)_P^1, \hat{f}(x)_P^2, \dots, \hat{f}(x)_P^t, \dots\}$  and telemetry data prediction set  $Q, \hat{f}(x)_Q = \{\hat{f}(x)_Q^1, \hat{f}(x)_Q^2, \dots, \hat{f}(x)_Q^t, \dots\}$  respectively. The prediction sequence  $\hat{Y}$  is obtained by integrating it with the weight matrix  $\zeta$ .

$$\hat{f}(x) = \begin{Bmatrix} \hat{f}(x)_P^1 & \hat{f}(x)_P^2 & \dots & \hat{f}(x)_P^t & \dots \\ \hat{f}(x)_Q^1 & \hat{f}(x)_Q^2 & \dots & \hat{f}(x)_Q^t & \dots \end{Bmatrix} \quad (3.9)$$

$R_t^*$  as some input at time  $t$ . The corresponding remote control instruction matrix  $D_t^*$  obtains the predicted value  $\hat{f}(x)_P^{t+s_t+1}$  at time  $t+s_t+1$  through the prediction model  $P$ .  $D_t^*$  obtains the predicted value  $\hat{f}(x)_Q^{t+s_t+1}$  at time  $t+s_t+1$  through the prediction model  $Q$ . If  $D_t^*$  corresponds to  $S2$  norm  $\|D_t^*\|_2$  belongs to the training set  $P$ , then:

$$\hat{f}(x)^{t+s_t+1} = \zeta^{t+s_t+1} \hat{f}(x)_P^{t+s_t+1} + (1 - \zeta^{t+s_t+1}) \hat{f}(x)_Q^{t+s_t+1} \quad (3.10)$$

If  $D_t^*$  corresponds to  $S2$  norm  $\|D_t^*\|_2$  belongs to training set  $Q$ .

**4. Experimental verification and result analysis.** The test data used in this paper are from the NASA Report published by NASA. It contains Acoustic Memory Activities and Cryptography (SMAP) and the Marten Stein Science Laboratory (MSL). Finally, the remote sensing data of the fusion LSTM model is analyzed, and the results are compared with those of the traditional LSTM model.

**4.1. Experimental Settings.** The training and calibration data include the telemetry data and the corresponding telemetry indication of each channel of 12 subsystems in SMAP and MSL. Ninety-four data sets in total. There were 47 groups (50%) with SMAP. There are 47 groups (50%) that are MSL. Each training set contains 11 remote control commands [13]. The number of remote sensing data sets varies from 300 to 4,000. The specific total number of training data, abnormal sequences, test telemetry channels and detection data are shown in Table 1.

**4.2. Comprehensive LSTM tests the anomaly detection algorithm in remote sensing data.** The abnormal fragments are compared with the actual fragments. Results, such as the exception detection component, are illustrated in Figure 4.1. As you can see from Figure 6, the exceptions found are essentially the same as those flagged [14]. Due to error redundancy  $q$ , the difference between the number of anomalies detected in this paper and the number of real anomalies is about  $\pm 100$ .

Table 4.1: Experimental data.

	SMAP	MSL	Total
Anomalous sequential total	72	38	109
The sum of dotted anomalous series	48(66.67%)	22(57.89%)	81(74.31%)
Background exception value total	24(33.33%)	16(42.11%)	28(25.69%)
Remote channel total	57	28	49
Total number of remote sensing data detected	447641	69489	517129

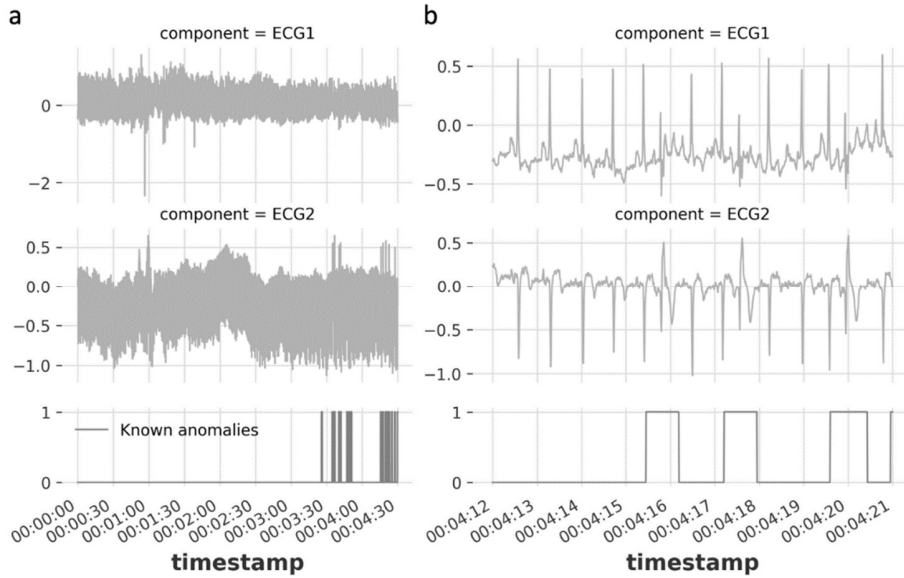


Fig. 4.1: Comparison between the actual anomalous sequence and anomalous detection results.

**4.3. Comparison of actual anomalous sequence and anomalous sequence detection results.** The ability of LSTM prediction is constantly improving, and the sensitivity to abnormal signals is also constantly improving. The change rule is evident for some remote sensing data, and remote control commands do not affect the remote sensing data. Therefore, anomaly detection in this case still needs further improvement [15]. The anomalies and background anomalies in two satellite remote sensing data types are examined based on ten experiments. FIG. 4.2 shows the performance comparison between an integrated LSTM-based anomaly detection method and an LSTM-based anomaly detection method in point anomaly detection and context anomaly detection for SMAP and MSL2 spacecraft.

A new method for anomaly detection of satellite remote sensing data using integrated LSTM is proposed. Compared with the LSTM algorithm, the single point anomaly detection rate of the SMAP algorithm is increased from 87.5% to 98.96%. The anomaly detection rate of SMAP applied to text content increased from 80.21% to 91.67%. The anomaly detection rate of MSL text was increased from 67.71% to 79.17%. A multimodal LSTM model based on the constraint model is proposed and applied to remote sensing data prediction. This can improve the recognition rate of abnormal events. Compared with the LSTM algorithm, this algorithm's overall point anomaly detection rate is increased from 73.96% to 81.25%. The abnormal detection rate of the whole environment increased from 75.00% to 84.38%.

**5. Conclusion.** This paper introduces an automatic control and management method for space vehicle test equipment. The monitoring and control of the test equipment are realized. This interface matches the

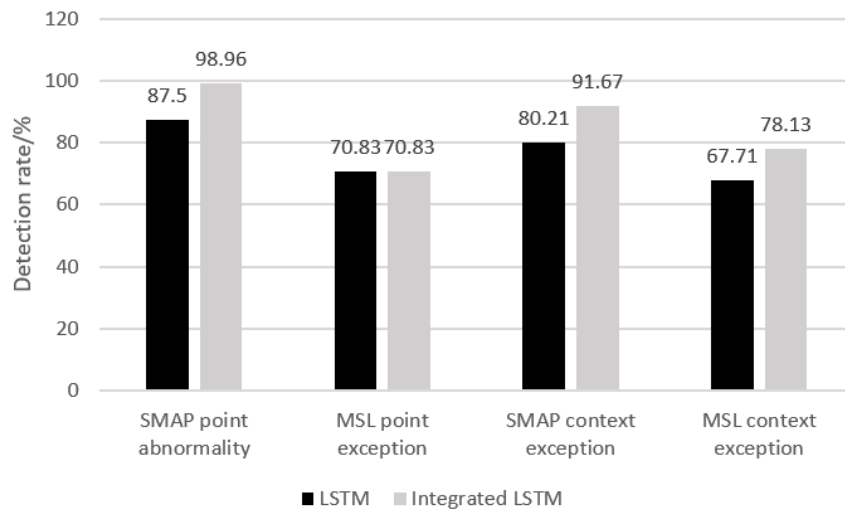


Fig. 4.2: Detection rates of different anomalies by two methods.

central control ATS. The system can automate and remotely manage the front-end test equipment of spacecraft. A spacecraft remote sensing measurement system is proposed, and experiments verify its effectiveness. An anomaly detection method for spacecraft telemetry data based on an integrated LSTM model is proposed. A dynamic threshold is established, and the difference between prediction and measurement is analyzed. The classification of remote measurement data is realized by mining the control mode of the remote command. This method improves the prediction level of remote sensing data by the LSTM model. The application of this system can effectively improve the management efficiency of front-end test equipment and promote the implementation of spacecraft automated tests.

**Acknowledgments.** The National Natural Science.

#### REFERENCES

- [1] Shi, H. B., Huang, D., Wang, L., Wu, M. Y., Xu, Y. C., Zeng, B. E., & Pang, C. (2021). An information integration approach to spacecraft fault diagnosis. *Enterprise information systems*, 15(8), 1128-1161.
- [2] Häseker, J. S., Aksteiner, N., & Ofenloch, A. (2021). From Behavioural to In-Depth Modelling of a Scalable Spacecraft Power System in SystemC-AMS using Continuous Integration Techniques. *Transactions of the Japan Society for Aeronautical and Space Sciences, Aerospace Technology Japan*, 19(5), 709-718.
- [3] Zheng, K., Zhang, Z., Chen, Y., & Wu, J. (2021). Blockchain adoption for information sharing: risk decision-making in spacecraft supply chain. *Enterprise Information Systems*, 15(8), 1070-1091.
- [4] Shen, B., Da Xu, L., Cai, H., Yu, H., Hu, P., Jiang, L., & Guo, J. (2022). Enhancing context-aware reactive planning for unexpected situations of on-orbit spacecraft. *IEEE Transactions on Aerospace and Electronic Systems*, 58(6), 4965-4983.
- [5] Zhang, C., Wu, J., Huang, Y., Jiang, Y., Dai, M. Z., & Wang, M. (2021). Constructive schemes to spacecraft attitude control with low communication frequency using sampled-data and encryption approaches. *Aircraft Engineering and Aerospace Technology*, 93(2), 267-274.
- [6] Lin, L. I., Li, Y. U. A. N., Li, W. A. N. G., Zheng, R., Yanpeng, W. U., & Xiaoyan, W. A. N. G. (2021). Recent advances in precision measurement & pointing control of spacecraft. *Chinese Journal of Aeronautics*, 34(10), 191-209.
- [7] Álvarez, J. M., & Roibás-Millán, E. (2021). Agile methodologies applied to Integrated Concurrent Engineering for spacecraft design. *Research in Engineering Design*, 32(4), 431-450.
- [8] Genova, A., & Petricca, F. (2021). Deep-space navigation with intersatellite radio tracking. *Journal of Guidance, Control, and Dynamics*, 44(5), 1068-1079.
- [9] Yin, Z., Cai, Y., Ren, Y., Wang, W., & Chen, X. (2020). A high precision attitude measurement method for spacecraft based on magnetically suspended rotor tilt modulation. *IEEE Sensors Journal*, 20(24), 14882-14891.
- [10] Schmidt, L., Gamper, E., Feuerle, T., & Stoll, E. (2020). Evaluation of the impact on European air traffic by uncontrolled reentering spacecraft with initial state uncertainty. *CEAS Aeronautical Journal*, 11(2), 401-416.

- [11] Zhu, W., Du, H., & Li, S. (2022). On event-triggered nonsmooth attitude tracking controller for a rigid spacecraft. *International Journal of Robust and Nonlinear Control*, 32(2), 900-916.
- [12] Pehrson, N. A., Ames, D. C., Smith, S. P., Magleby, S. P., & Arya, M. (2020). Self-deployable, self-stiffening, and retractable origami-based arrays for spacecraft. *AIAA Journal*, 58(7), 3221-3228.
- [13] Mote, M., Egerstedt, M., Feron, E., Bylard, A., & Pavone, M. (2020). Collision-inclusive trajectory optimization for free-flying spacecraft. *Journal of Guidance, Control, and Dynamics*, 43(7), 1247-1258.
- [14] Hovell, K., & Ulrich, S. (2021). Deep reinforcement learning for spacecraft proximity operations guidance. *Journal of spacecraft and rockets*, 58(2), 254-264.
- [15] Yu, X., Zhu, Y., Qiao, J., & Guo, L. (2021). Antidisturbance controllability analysis and enhanced antidisturbance controller design with application to flexible spacecraft. *IEEE Transactions on Aerospace and Electronic Systems*, 57(5), 3393-3404.

*Edited by:* Zhigao Zheng

*Special Issue on:* Graph Powered Big Aerospace Data Processing

*Received:* Jul 14, 2023

*Accepted:* Aug 25, 2023



## RESEARCH ON SPACE IMAGE FAST CLASSIFICATION BASED ON BIG DATA

**Abstract.** In order to improve the accuracy and effect of space image classification, the author proposes a space image classification method based on Big data analysis, aiming at the shortcomings of low accuracy and long time of current image classification. First, analyze the current research progress of image classification, find out the shortcomings of different classification methods, then collect aerospace images, preprocess the images, and use big data analysis technology to establish image classifiers, image classification was performed using an image classifier, and finally simulation experiments were conducted with other methods for image classification. The results indicate that: The average classification time of this method for aerospace images is 3.5 minutes, which saves 14 minutes and 29 minutes compared to traditional method 1 and traditional method 2, respectively. This indicates that this method has the shortest image classification time and improves the classification efficiency of aerospace images. This method has been proven to have high accuracy in image classification, the shortest classification time, and significant advantages compared to other image classification methods.

**Key words:** Big data analysis; Image classification; Convolutional neural network; Classification accuracy; Classification effect

**1. Introduction.** In recent years, intelligent spacecraft have increasingly attracted the attention of aerospace practitioners. Image recognition is one of the main conditions for spacecraft intelligence, and it is also an important research topic in fields such as computer vision, machine learning, and pattern recognition. With the rapid development of computing technology and image sensors, image acquisition methods have been expanded and the field of vision has been promoted. More and more devices have the ability to obtain images, sparking a wave of device intelligence [1]. However, limited by the computing power of flight controllers in current spacecraft, popular deep learning models typically require massive amounts of data and storage space, as well as a large amount of computing resources for long-term model training. If applied in spacecraft, specialized computing equipment is required, which increases the cost and takeoff quality of the spacecraft. At the same time, the space scene is constantly changing over time. By solving the problem of fast online training, spacecraft can have the ability to autonomously recognize targets and have better adaptability to the space environment. Therefore, it is necessary to propose a fast image classification method that can be flexibly adjusted according to practical applications at the software level. Big data analysis methods have been widely used in the field of natural image classification, and detection, registration, generation and other technologies are also gradually applied in image classification [2]. The big data analysis method is used for multi-level processing of information, simulating human brain thinking mode. Different spaces correspond to different layers of features, with different semantic information, hierarchical feature structure of classification, and high classification ability. In order to obtain ideal image classification results, a space image classification method based on Big data analysis is proposed, and its performance is analyzed.

**2. References.** Computer vision is the basic way to achieve machine intelligence, and its basic carrier is image. The rapid development of computing technology and image sensors has provided great convenience for the acquisition and sharing of images, and the scale of image data is showing explosive growth. From obtaining planetary optical images to deep space exploration, from video surveillance to early warning reconnaissance, image and image processing technologies are ubiquitous [3]. Image classification is one of the hot fields in computer vision, machine learning, and pattern recognition, and has also been widely applied in the aerospace field. Image classification consists of two parts: classifier learning and testing. During learning, use the feature information extracted from the image to train the classifier and form classification rules. During testing, each identified sample is described as a set of feature vectors, and the classifier determines the category of the feature vectors based on the learned classification rules. Therefore, image classification can be seen as the process of completing the mapping from the "feature space" to the "category space". After years of development, image

classification has proposed many effective learning models, including BP networks, support vector machines, and deep learning networks [4]. The BP network uses a sample set to train the network, thereby obtaining the rules within it. It can fit any function with any accuracy, and the learning rules are simple and easy to implement, which has been widely applied. However, the neural network also has obvious shortcomings, such as the training effect is limited by the network size, and the parameter setting theory is not perfect, which leads to the full Rate of convergence of the neural network, and the classification performance is affected by the fitting effect [5].

Due to the difficulty in overcoming these issues, the popularity of related research has gradually decreased, and researchers have turned their attention to support vector machines. Support vector machines use sum functions to classify nonlinear problems through mapping. The advantage of support vector machines is that they have a complete theoretical foundation, are suitable for small sample learning problems, and have high learning efficiency. Therefore, they have been widely applied since their introduction [6]. The above two classification models are both Supervised learning models. The disadvantage is that the premise of learning is that experts need to provide category information of samples. The deep learning method is a redevelopment of traditional neural network methods, and has achieved many remarkable achievements in fields such as speech recognition and artificial intelligence. It is currently one of the cutting-edge research content in the field of image classification. The deep learning method believes that the training difficulty of multi-layer neural networks can be effectively overcome through unsupervised methods. Deep learning can be achieved by learning deep nonlinear neural networks, mapping low-level features to higher-level forms of features, and learning features with hierarchical structures from them. This not only preserves the advantage of traditional neural networks being able to approximate complex functions with arbitrary accuracy, but also solves the problem of parameter tuning and overfitting in traditional neural network methods, significantly improving the accuracy of image classification, it has received widespread attention from the academic community since its inception, and various types of research and applications have emerged endlessly. But the drawbacks of deep learning are also very obvious. Firstly, deep learning methods weaken image feature extraction, resulting in the learning process requiring massive amounts of data to obtain satisfactory results, resulting in low learning efficiency; Secondly, the learning process requires a large amount of computing resources, which has high time and Space complexity, and is difficult to deploy on the resource constrained rocket borne computer. It usually requires additional dedicated hardware to complete the classification task, increasing the flight cost [7].

Zhengwen Li proposes a new image classification method based on the cleaning of inaccurate image data. The effectiveness and effectiveness of this method have been demonstrated through testing real cat and dog images. In the process of studying the relationship between the proportion of incorrectly labeled images in the dataset and classification accuracy, we found that deeper neural networks have a certain degree of robustness to erroneous images in the dataset. However, when there is a high proportion of tag noise images in the image data set [8]. On this basis, Wang .J proposed a data integration scheme for the enterprise Human resource management system based on Big data. Firstly, a unified standard for EIS data integration was established and EIS data was classified. Based on the relationship between data, data types were defined and the data in the employee information system was classified. After classification, inconsistent and duplicate data was eliminated to reduce the interference of invalid data and improve the efficiency of data integration [9]. Based on existing power supply area data, Li and Z have defined four indicators: minimum negative line loss rate, maximum positive line loss rate, minimum power factor, and maximum number of negative line loss values. On this basis, combined with the four indexes before, after, after and after error correction, and combined with the decision tree algorithm, the Discriminative model of the error correction area after error correction is constructed. This identification model was used to identify incorrect wiring areas in all areas of the power company, and on-site inspections were conducted on areas identified as incorrect wiring by the model [10].

Based on the current research, this paper proposes a space image classification method based on Big data analysis. First, analyze the current research progress of image classification, find out the shortcomings of different classification methods, then collect aerospace images, preprocess the images, and use Big data analysis technology to establish image classifiers, and use image classifiers to classify images. Finally, carry out simulation experiments with other methods for image classification.



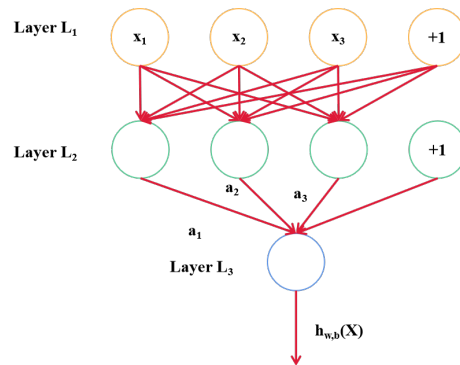


Fig. 3.1: Structural diagram of the artificial neural network model.

**3. Space image classification method based on Big data analysis.** Convolutional neural network (CNN) is a Big data analysis technology, which simulates the information processing process of human brain, and uses Convolutional neural network algorithm to classify images, the neural network is composed of several neurons, which generate high-level abstract features through low-level single features, and is applied to aerospace image classification.

**3.1. Composition of Convolutional neural network.** Convolutional neural network is evolved from the basic model of artificial neural network. The artificial neural network model is composed of multiple nodes, which form a network according to a specific connection mode. The data flows into the network from the input node, undergoes a series of calculations and transfers, and the results are obtained at the output node. The basic unit of this model is neurons. Each neuron receives input data, processes the input through a certain Activation function, and then transmits the processed results to the next layer of neurons. This connection method forms a multi-layer neural network, with each layer having different feature representation capabilities. The higher the level, the more accurate the semantic representation of the features. The structure diagram of the artificial neural network model is shown in Figure 3.1. As shown in Figure 3.1, the artificial neural network model connects various data nodes and calculates the results after the data flows into the network. Convolutional neural network introduces convolution layer and pooling layer on the basis of artificial neural network, which is the main difference between it and traditional neural network. The convolutional layer utilizes convolutional operations to extract local features from input data, effectively capturing spatial relationships in the data through shared weights and local connections [11]. The output of the convolutional layer is transmitted to the pooling layer, which is used to reduce the size of the feature map and retain the main feature information, thereby reducing computational complexity and enhancing the robustness of the model. Convolutional neural network gradually extracts advanced feature representation of input data by stacking multiple convolution layers and pooling layers. Finally, the extracted features are mapped to the output nodes through a fully connected layer to obtain the final classification or regression results [12]. The multi-layer structure of Convolutional neural network enables it to automatically learn the feature representation of input data, and has certain invariance to translation, scale, deformation and other transformations. This makes Convolutional neural network perform well in computer vision and image processing tasks, such as image classification, object detection and image generation. In addition, Convolutional neural network can also be applied to tasks in other fields such as text data and time series data. In a word, Convolutional neural network constructs a multi-level neural network model with hierarchical feature representation ability by introducing convolutional layer and pooling layer to adapt to data analysis and processing requirements in different fields.

(1) Convolutional layer. The original input matrix utilizes a convolutional kernel sampler to generate convolutional layers.

(2) Downsampling layer. Implement pooling computing based on convolutional layers and generate downsampling layers.

(3) Fully connected layer. By iteratively constructing multiple intersecting sampling layers and generating

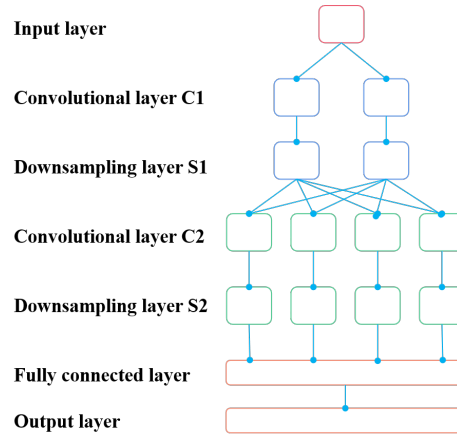


Fig. 3.2: Structural diagram of the artificial neural network model.

fully connected layers with multiple convolutional layers, CNN construction is completed. The CNN structure is shown in Figure 3.2.

**3.2. CNN training process.** By convolving the sample matrix in the first layer,  $y$  is obtained through multi-level transformation. Assuming  $y$  represents the expected output of the sample, the error between the two is described by  $E$  [13]. When performing backpropagation, fine-tuning the convolutional kernel matrix is based on the principle of minimizing error. Assuming the output of the  $y$ -th neuron in layer  $v$ , there is:

$$c_j^v = \sum_k m_{jk}^v e_k^{v-1} + d_j^v \quad (3.1)$$

Among them, the correction term, the output of the  $k$ -th neuron in layer  $v - 1$ , and the weight of the connection between the  $j$ -th and  $k$ -th neurons in layer  $v$  are described by  $d_j^v$ ,  $e_k^{v-1}$ ,  $m_{jk}^v$ , respectively. The output of the  $y$ -th neuron in layer  $v$  is described by  $c_j^v$ . The formula for calculating the output  $e_k^{v-1}$  of the  $k$ -th neuron in layer  $v - 1$  is shown in formula (3.2):

$$e_k^{v-1} = \delta(c_j^{v-2}) \quad (3.2)$$

Where: The Activation function is described by  $\delta$ .

The Error function is calculated according to the error between the calculated value and the expected value. The specific calculation process is shown in Formula (3.3):

$$\theta = f(y, y') \quad (3.3)$$

Among them, the Error function is described by  $\theta$ , and the quadratic Algebraic function is described by  $f$ . The error calculation formula for the  $j$ -th neuron of layer  $v$  is shown in formula (3.4):

$$\vartheta_j^v = \frac{\partial \theta}{\partial c_j^v} \quad (3.4)$$

The formula for calculating the last layer error of CNN is shown in formula (3.5):

$$\vartheta^v = \nabla_{e^v} \theta \Theta \delta'(c^v) \quad (3.5)$$

Among them, the gradient value of the last layer, the product operator, and the output of the  $v$  layer are described by  $\nabla_{e^v}$ ,  $\theta$ ,  $\Theta$ ,  $c^v$ , respectively.

The error calculation formula for other layers is shown in (3.6):

$$\vartheta^v = \left( (m^{v+1})^T \vartheta^{v+1} \right) \Theta \delta' (c^v) \tag{3.6}$$

The weight of  $v + 1$  layer is described by  $m^{v+1}$ , the error of  $v + 1$  layer is described by  $\vartheta^{v+1}$ , and the function is described by  $T$ .

The weight gradient calculation formula is shown in (3.7):

$$\frac{\partial \theta}{\partial m_{jk}^v} = \vartheta_j^v e_k^{v-1} \tag{3.7}$$

The formula for calculating the bias gradient is:

$$\frac{\partial \theta}{\partial d_j^v} = \vartheta_j^v \tag{3.8}$$

The update formula for convolutional kernels is to use gradient descent, and the update formula for convolutional kernels is as follows:

$$m^v = m^v - \eta \sum_x \theta^{x,v} (e^{x,v-1})^T \tag{3.9}$$

$$d^v = d^v - \eta \sum_x \theta^{x,v} \tag{3.10}$$

Among them,  $m^v, d^v$  represents the weight and correction term of layer  $v$ .

**3.3. Space image classification process based on Big data analysis.** The aerospace image classification process based on Big data analysis is to achieve aerospace image classification by constructing a five layer Convolutional neural network. The following are the detailed steps:

(1) Data collection and preprocessing: Collect a large number of aerospace image datasets, including images of different categories and scenarios. Preprocess the collected image data, including image denoising, size normalization, brightness adjustment, etc., to improve the accuracy and robustness of subsequent classification [14].

(2) Build a Convolutional neural network (CNN) model:

Design a five layer Convolutional neural network structure. It usually includes convolution layer, Activation function layer, pooling layer and full connection layer.

Select the appropriate number and size of convolutional and pooling layers based on specific tasks and data characteristics, and decide whether to add batch normalization and regularization layers. For each convolution layer and fully connected layer, an appropriate Activation function (such as ReLU) is defined to introduce non-linear transformation. Set the number of neurons in the output layer to be equal to the number of classification categories of aerospace images, and use an appropriate Activation function (such as Softmax) for multi category classification.

(3) Parameter adjustment and model training: initialize the weight and bias parameters of the Convolutional neural network. Divide the dataset into a training set and a validation set, and use the training set for model training. The network parameters are adjusted by Backpropagation and optimizer (such as random gradient descent) to make the model gradually adapt to the training data and minimize the Loss function. In the training process, monitor the classification accuracy and Loss function values on the validation set, and adjust the model structure and super parameters (such as Learning rate and batch size) as needed.

(4) Model evaluation and tuning: use independent test sets to evaluate the performance of the trained Convolutional neural network model. Analyze the classification accuracy, accuracy, recall, and F1 values of the model in different categories to evaluate its classification performance. Based on the evaluation results, adjust the network structure, optimize algorithms, and hyperparameters to further improve the performance of the model [15].

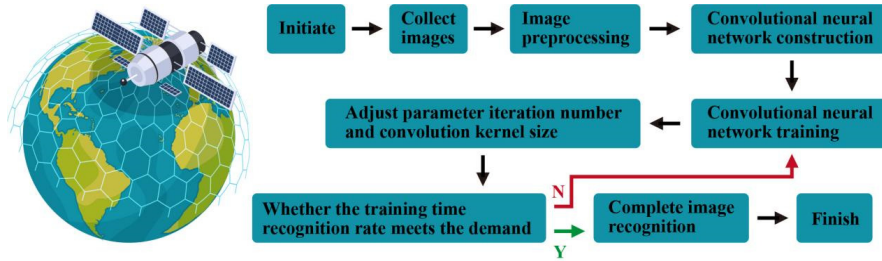


Fig. 3.3: Massive image classification process based on big data analysis.

(5) Space image classification: use the trained Convolutional neural network model to classify new space images. Input the images to be classified into the model and calculate the probability distribution of each category through forward propagation. Assign aerospace images to corresponding classification categories based on the highest probability category or set threshold.

Through the above detailed steps, the aerospace image classification process based on Big data analysis uses five layer Convolutional neural network to classify aerospace images. This process can improve the accuracy and efficiency of image classification, and provide strong support for aerospace image analysis and application.

The process of aerospace image classification based on big data analysis is as follows: first collect aerospace image data, which is divided into test data and training data; Implement preliminary data preprocessing for aerospace images, decolor the aerospace image, convert it to Grayscale, extract some features of the aerospace image, and implement normalization processing to  $30 \times 30$ ; The convolutional kernel is  $5 \times 5$ . A fully connected layer, two downsampling layers, and two convolutional layers form the CNN model, the lower sampling layer uses non repeated pooling with a scale of 2. Six feature maps constitute the first layer of convolution, and 16 feature maps constitute the second layer of convolution [16]. The output categories are multi categories, and the sigmoid function is the Activation function; The Convolutional neural network is used to change the model structure, adjust the iteration number, convolution core size and other parameters, and determine whether the training time and classification rate meet the requirements, when the requirements are met, the classification of aerospace images can be completed. If not, Convolutional neural network training needs to be conducted again until the requirements are met. Therefore, the aerospace image classification process based on Big data analysis is shown in Figure 3.3. Aerospace image data preprocessing process: Due to the impact of external environment, the quality of aerospace images decreases, and preprocessing of aerospace images is required. The aerospace image is divided into several sub images, and 8 sub images are set. The discrete cosine wave transformation calculation process of the sub images is shown in formula (3.11):

$$F(u, v) = \frac{1}{4} C(u) C(v) \sum_{m=0}^7 \sum_{n=0}^7 f(m, n) \cos \frac{(2m+1)u\pi}{16} \cos \frac{(2n+1)v\pi}{16} \tag{3.11}$$

Among them, the sub images are described by  $f(m, n)$ ,  $0 \leq u \leq 7, 0 \leq v \leq 7$ ,

$$C(u) = C(v) = \begin{cases} \frac{1}{\sqrt{2}}, & u = v \\ 0, & \text{otherwise} \end{cases} \tag{3.12}$$

The inverse transformation using discrete cosine transform is shown in formula (12):

$$F(x, y) = \frac{1}{4} \sum_{u=0}^7 \sum_{v=0}^7 C(u) C(v) F(u, v) \cos \frac{(2m+1)u\pi}{16} \cos \frac{(2n+1)v\pi}{16} \tag{3.13}$$

In the case of a small number of samples, the data preprocessing process is as follows:

- (1) Enhance the dataset using mirror symmetry method.
- (2) Implement background segmentation on scratch datasets with prominent features.
- (3) Implement principal component analysis for dimensionality reduction of transformed aerospace images.

#### 4. Simulation experiments.

**4.1. Experimental subjects.** To verify the effectiveness of our method in aerospace image classification, we conducted simulation experiments in the Matlab R2013c experimental environment of the Windows 7 operating system. We selected 15000 images as experimental samples, of which 10000 were used as training data and the remaining 5000 were used as testing data. We compare the method proposed in this paper with the image classification methods of Traditional Method 1 and Traditional Method 2, and evaluate their differences in image classification accuracy, performance under different iterations, classification error curve, and image classification time. During the experiment, we used the method proposed in this paper and traditional methods 1 and 2 to classify the training data into images, and evaluated their classification accuracy on the test data. We also compared their performance under different iterations and the changes in classification error curves. In addition, we also recorded the time required for image classification to evaluate the computational efficiency of different methods [17]. Through the above comparative experiments, we can comprehensively evaluate the performance of our method in aerospace image classification and compare it with traditional methods to verify the advantages and effectiveness of our method. The experimental results will provide detailed information about the classification accuracy, Rate of convergence, calculation efficiency and other aspects of different methods, thus providing a reference for the practical application of aerospace image classification tasks.

**4.2. Image classification accuracy.** We classified 500 images using traditional methods 1, 2, and our method, and compared their classification accuracy. Figure 4.1 was drawn based on experimental results, demonstrating the classification accuracy of the three methods. From the experimental results in Figure 4.1, it can be concluded that the proposed method and image average classification accuracy are significantly higher than traditional methods 1 and 2, exhibiting better image classification results. This means that in the aerospace image classification task, our method can more accurately classify images into the correct categories, and has higher classification accuracy compared to traditional methods 1 and 2 [18]. This result may be due to the Convolutional neural network model used in this method, which can better extract image features and capture the correlation between images. In contrast, traditional methods 1 and 2 may have certain limitations in feature extraction and pattern recognition, resulting in lower classification accuracy. Therefore, based on the experimental results in Figure 4.1, it can be concluded that in the aerospace image classification task, our method can achieve higher image classification accuracy compared to traditional methods 1 and 2, providing better results and performance for aerospace image classification.

**4.3. Classification Error Curve for Different Iterations.** Considering the impact of the relationship between the number of samples in the training set and the training error on the image classification performance, the number of iterations is 30 and 60, respectively. The classification error curves for different iterations are described in Figure 4.2. As shown in Figure 4.2, the training error of this method is inversely proportional to the number of iterations and the number of samples in the training set, as the number of iterations and the number of training samples increase, the training error also decreases. In Figure 4.2(b), the minimum training error of this method is 0.02, the minimum training error of traditional method 1 is 0.08, and the minimum training error of traditional method 2 is 0.11, indicating that this training error is the smallest.

**4.4. Space image classification time.** Compare the classification time of aerospace images using three methods, and the specific results are shown in Figure 4.3. From the results of Figure 4.3, it can be seen that the average classification time of this method for aerospace images is 3.5 minutes, which is 14 minutes and 29 minutes less than traditional method 1 and traditional method 2, respectively. This indicates that this method has the shortest image classification time and improves the classification efficiency of aerospace images [19, 20, 21, 22, 23].

**4.5. Conclusion.** This paper aims to design a space image classification method based on Big data analysis, and test it through simulation experiments. The experimental results show that the proposed method

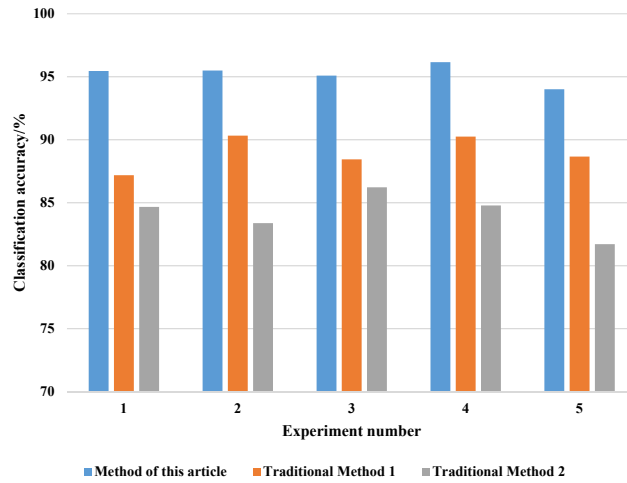
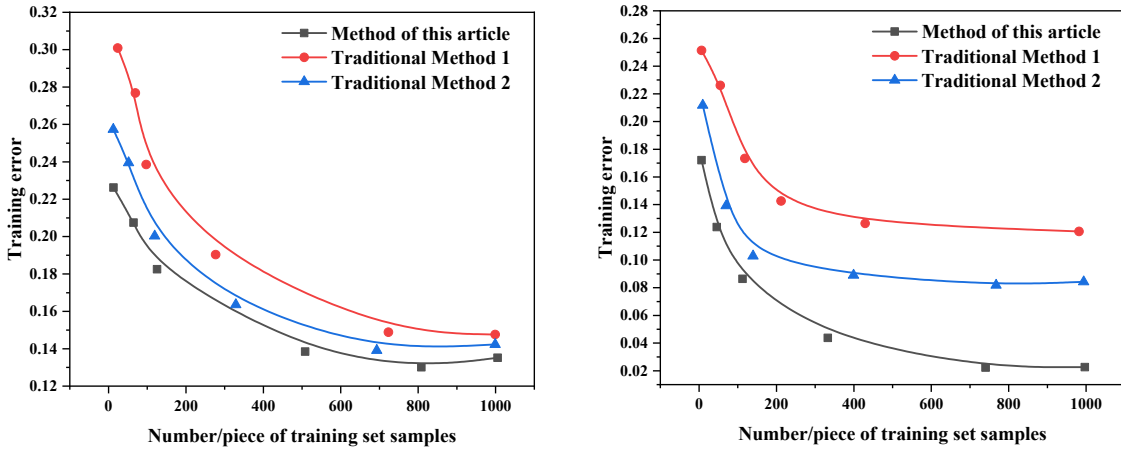


Fig. 4.1: Classification results of the three methods.



(A) Classification error curve when the number of iterations is 30 times

(B) Classification error curve when the number of iterations is 60

Fig. 4.2: Classification results of the three methods.

exhibits the following advantages in aerospace image classification tasks: high accuracy, small training error, short image classification time, and good image classification performance. This indicates that the method proposed in this paper can effectively improve the efficiency of aerospace image classification and provide a strong theoretical basis for subsequent image processing work. However, there are also some shortcomings in this study, including the following aspects: limited time and energy: due to time and energy limitations, this study may not be able to cover all possible situations and details. Therefore, the image classification effects in certain specific scenarios or special circumstances may not have been fully explored. In order to further improve the method of this article, future research can consider the following aspects:

1. Integration of advanced science and technology: With the continuous development of science and technology, more advanced technologies and methods can be integrated into the methods in this paper, such as deep learning, Transfer learning, attention mechanism, etc. These technologies can further improve the speed and accuracy of aerospace image classification.

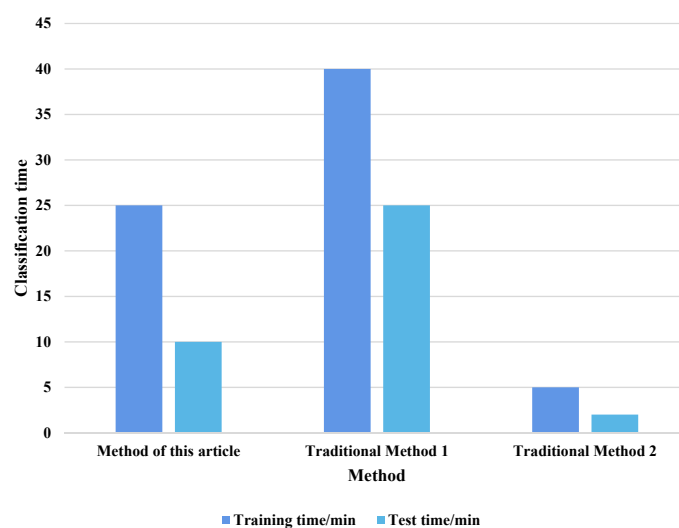


Fig. 4.3: Classification results of the three methods.

2. Data augmentation technology: Through data augmentation technology, the training dataset can be expanded, the diversity and quantity of samples can be increased, and the generalization ability and robustness of the model can be improved.

3. Model optimization and parameter adjustment: further optimize the Convolutional neural network model structure, and further improve the image classification performance through appropriate parameter adjustment methods.

In summary, although there are some shortcomings in this study, by integrating more advanced scientific and technological knowledge into this method, and further optimizing and improving it, aerospace image classification can be achieved more quickly, providing better support for image processing work in the aerospace field.

**Acknowledgments.** The National Natural Science Foundation of China (41601394).

#### REFERENCES

- [1] Yang, J. . (2021). Folk song computer big data classification and analysis research based on national style characteristics. *Journal of Physics: Conference Series*, 1744(3), 032117 (5pp).
- [2] Ding, J. , Zhao, G. , Huang, J. . (2022). Research on block image texture retrieval method based on depth hash. *Journal of Interconnection Networks*,58(1),456-457.
- [3] Chen, J. , Li, H. , Jiang, H. , Liang, T. . (2021). Research on anthrax toxin analysis based on big data mining technology. *Journal of Physics: Conference Series*, 1774(1), 012005 (6pp).
- [4] Lin, C. Y. , Xi, Z. , Gao, C. , Tsai, S. B. . (2021). Research on the training model of e-commerce professionals based on big data analysis. *Wireless Communications and Mobile Computing*,157(1),96-99.
- [5] Ji, L. (2021). Research on fast de-duplication of text backup information in library database based on big data. *International Journal of Information and Communication Technology*, 19(1), 76.
- [6] Xing, Y. , Wang, X. , Qiu, C. , Li, Y. , He, W. . (2022). Research on opinion polarization by big data analytics capabilities in online social networks. *Technology in Society*, 68(8),59-60.
- [7] Zong, X. , Zhang, C. , Wu, D. . (2022). Research on data mining of sports wearable intelligent devices based on big data analysis. *Discrete Dynamics in Nature and Society*,187(33),749-759.
- [8] ZhengwenLI, WenjuDU, NiniRAO. (2022). Research on classification method based on inaccurate image dataset cleaning. *Journal of Signal Processing*, 38(7), 1547-1554.
- [9] Wang, J. , Li, D. , Wang, Z. , Wan, T. . (2021). Research on enterprise employee information system based on big data analysis. *Journal of Physics: Conference Series*, 1748(3), 032025 (6pp).
- [10] Li, Z. , Yu, L. , Liu, J. , Su, L. , Tan, H. , Zhao, J. , et al. (2021). Modeling and algorithm research on identification of wrong wiring power supply region based on classification analysis. *Journal of Physics: Conference Series*, 2068(1), 012036.

- [11] Wang, H. , Zuo, H. , Liu, Z. , Fei, H. , Liu, Y. , Jiang, Z. . (2021). Research on 3d reconstruction method of wear particle dynamic image based on multi contour space mapping. IOP Conference Series: Materials Science and Engineering, 1207(1), 012017.
- [12] Chebbi, I. , Mellouli, N. , Farah, I. R. , Lamolle, M. . (2021). Big remote sensing image classification based on deep learning extraction features and distributed spark frameworks. Big Data and Cognitive Computing, 5(2), 21.
- [13] Qi, B. . (2021). Hyperspectral image database query based on big data analysis technology. E3S Web of Conferences, 275(1), 03018.
- [14] Huang, J. C. , Ko, P. C. , Fong, C. M. , Lai, S. M. , Chen, H. H. , Hsieh, C. T. . (2021). Statistical modeling and simulation of online shopping customer loyalty based on machine learning and big data analysis. Security and Communication Networks, 748(1), 45-52.
- [15] Sun, F. , Zhang, J. . (2021). Research on grammar checking system using computer big data and convolutional neural network constructing classification model. Journal of Physics Conference Series, 1952(4), 042097.
- [16] Jiang, H. , Chen, K. , Ge, Q. , Wang, Y. , Xu, J. , Li, C. . (2021). Fault diagnosis of power iot system based on improved q-kpca-rf using message data. IEEE internet of things journal, 98(8-11), 17-23.
- [17] Rogério Luís de C. Costa, José Moreira, Pintor, P. , Santos, V. D. , Sérgio Lifschitz. (2021). A survey on data-driven performance tuning for big data analytics platforms. Big Data Research, 66(8), 100206.
- [18] Punia, S. K. , Kumar, M. , Stephan, T. , Deverajan, G. G. , Patan, R. . (2021). Performance analysis of machine learning algorithms for big data classification: ml and ai-based algorithms for big data analysis. International Journal of E-Health and Medical Communications (IJEHMC), 12(1), 96.
- [19] Fu, C. , Tu, X. , Huang, A. . (2021). Identification and characterization of production-living-ecological space in a central urban area based on poi data: a case study for wuhan, china. Sustainability, 34(14), 58-63.
- [20] Liu, L. , Tsai, S. B. . (2021). Intelligent recognition and teaching of english fuzzy texts based on fuzzy computing and big data. Wireless Communications and Mobile Computing, 2021(1), 1-10.
- [21] Guo, L. , Wang, M. , Lin, Y. . (2021). Electromagnetic environment portrait based on big data mining. Wireless Communications and Mobile Computing, 2021(3), 1-13.
- [22] Li, Y. , Xia, C. , Zhuang, Y. , Liao, B. , Huang, S. . (2021). Research on residual current warning technology based on big data analysis. Journal of Physics: Conference Series, 2108(1), 012012.
- [23] Koulali, R. , Zaidani, H. , Zaim, M. . (2021). Image classification approach using machine learning and an industrial hadoop based data pipeline. Big Data Research, 24(9), 100184.

*Edited by:* Zhigao Zheng

*Special Issue on:* Graph Powered Big Aerospace Data Processing

*Received:* Jul 24, 2023

*Accepted:* Aug 25, 2023





## APPLICATION OF DATA VISUALIZATION INTERACTION TECHNOLOGY IN AEROSPACE DATA PROCESSING

**Abstract.** A visualization and interactive network topology model are studied based on real-time features generated during spaceflight. Start by establishing a consistent set of data and logical interaction interfaces. This paper presents a method of scenario model construction and application programming based on virtual reality technology. The scene elements are extracted into two types of primitives, namely logical type and simulated object type. This provides a unified architecture for the editing and processing of graphic elements. This system can realize the automatic creation of the scene. Then the point cloud data obtained by sparse reconstruction of SFM is reconstructed to the Poisson surface. You get a dense, uniform grid. Experiments show that the proposed algorithm can realize the 3D reconstruction of non-cooperative objects. The spatial feature points obtained in the spatial positioning of non-cooperative objects can provide necessary technical support for its orbit positioning. The model can quickly generate new model scenario pages according to the characteristics of the task. This method changes the display mode, which can only be static or limited dynamic before. It has also improved the efficiency of space mission preparation.

**Key words:** Space mission; Pixel; Visual interaction; Topology modeling; Three-dimensional reconstruction; Motion recovery structure (SFM)

**1. Introduction.** China's aviation technology innovation has grown by leaps and bounds to adapt to the fierce competition in space technology. China has completed several important space science and technology projects. Overall, it shows the development trend of multi-region, multi-model, multi-region and intensive firing. This has led to higher real-time demands across disciplines. It includes fast preparation, binding, switching, multiple presentations, and dynamic command and support scene construction. Traditional commercial applications such as space command and surveillance are a "waterfall" development process. Software requirements are based on user requirements. Then write a software requirements document. Then the software is designed, coded, tested and maintained. Due to the significant differences between the various types of spacecrafts, there is no single standard. Business systems software tends to be highly personalized. It is challenging to realize the repeated utilization of several tasks. Due to the lack of reusable software and efficient auxiliary methods, the development efficiency of the system is very low, and it isn't easy to meet the new development requirements.

Currently, the widely used network model construction technology of visualization software can be roughly divided into two categories: the personalized design based on interactive graphics. Put the image class into the program. However, this technology is more complicated, and the development time is relatively long. The number of codes increases geometrically. This is unfavorable to the expansion and maintenance of the software. One is a single development based on QtUiloader. This wraps the chart as a dynamic library that frames can access. It uses the "skeleton + plug-in" function in hot-swap mode. It reduces the difficulty of research and development and saves time. This approach reduces the maintenance complexity of the system. However, the industry has not yet established a universal 3D visual application model. Existing model-building methods are rarely applied. It only works for his business activities. This method has poor support for new devices, tasks, and models.

Construction from Motion (SFM) is a 3D reconstruction technique based on multi-frame images. Many foreign experts on this issue have done a lot of work. Literature [1] uses SFM technology for 3D reconstruction and develops a 3D automatic reconstruction system. Beam adjustment (BA) is an essential algorithm for SFM optimization. The point cloud data was obtained by minimizing the effect of repeated projections by the LM (Levenberg-Marquardt) method. Aiming at the problem that the current large amount of data is not suitable for large-scale reconstruction; some scholars intend to introduce a method based on sparse beam correction [2]. It reduces the size and complexity of operations. The dense point cloud is obtained using CMVS/PMVS (Clustering views Multi-View Stereo / Patch-based Multi-View Stereo) algorithm based on sparse point cloud

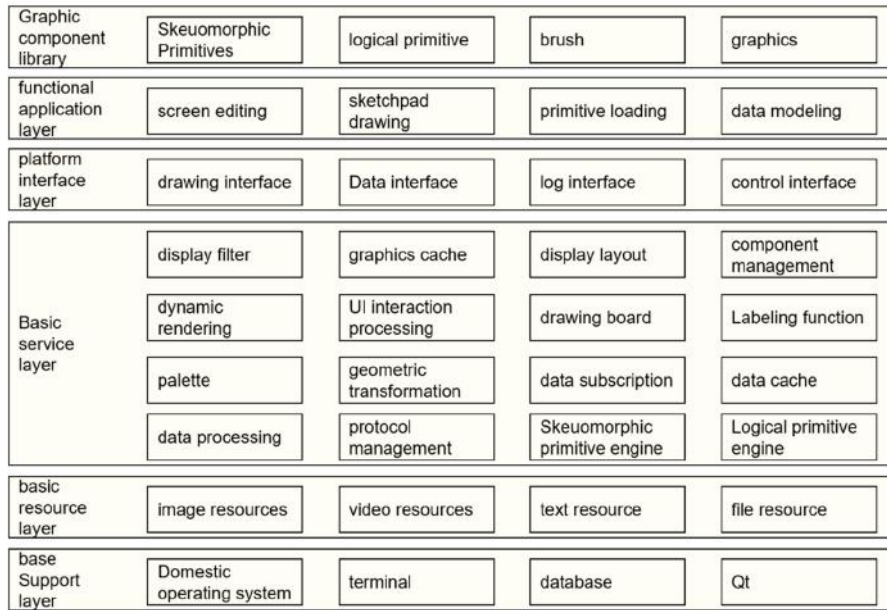


Fig. 2.1: Visual topology modeling system framework.

reconstruction. The final 3D reconstruction results are satisfactory [3]. There are also many domestic research results. Literature [4] has conducted in-depth research on the fast calibration, fast image matching and fast and accurate basic matrix calculation required by the system in the 3D reconstruction of space non-cooperative targets. Literature [5] uses the imaging timing of continuous images as iterative prior information and uses improved SFM for 3D reconstruction. Literature [6] researched 3D reconstruction for CMVS/PMVS algorithm. Experimental results prove that this method is very effective in outdoor reconstruction. Based on the 3D reconstruction algorithm based on the brightness and darkness of a single frame image to restore the shape, Literature [7] proposed an infrared-self-shadow reconstruction (IR-SFS) method using external radiation and its radiation for 3D reconstruction. The overall effect has been significantly improved.

This paper presents a method of constructing a network model based on visual interaction. Develop interactive data and logic interfaces. This paper presents a method of scenario model construction and application programming based on virtual reality technology. The scene elements are extracted into two graph units, namely logical and analog shape graph units, and the corresponding algorithms are constructed. The system has the function of judging and processing the bundling parameters. You can quickly create scene pages like decision trees, flow charts, relationship diagrams, dynamic pipeline diagrams, device diagrams, etc. The new number, task scene drags, visual topology modeling, graphics drawing, interactive display, and other functions are realized. The SFM algorithm is applied to the sparse point cloud reconstruction of non-cooperative objects. The CMVS/PMVS method reconstructs the dense point cloud data. At the same time, this project will also use Poisson reconstruction technology to improve the existing defects, such as non-uniform distribution and voids in the dense point cloud. Users can get a clear, intuitive view of non-cooperative targets. It realizes the function of space observation for ordinary users.

## 2. Development of network model construction system for visual interaction.

**2.1. Network model structure of visual interaction.** The visual architecture of "platform + plug-in" and "common + feature" is used to model the network topology. A set of web page editing systems is designed based on QtUiLoader. According to the different functions, the visual topology modeling system is divided into basic support layer, basic resource layer, basic service layer, platform interface layer, element component library and function application layer. A specific architecture diagram is shown in Figure 2.1.

The foundation support layer is the primary operating environment and supporting condition of the whole system. The focus is the self-developed operating system, hardware terminal, lightweight database, lightweight database, and Qt development platform.

The primary resource layer is the base load resource of the system. It is a vital resource investment. It provides image, video, and file resources for the system, such as device connection diagrams, cloud diagrams, geographic information resources, video monitoring resources, task introduction, personnel information, and task documents.

The underlying service layer is a platform-level function that implements the entire system. It is a general resource extraction for each component. Its content mainly includes a display filtering service, graphics cache service, display layout service, component management service, dynamic rendering service, UI interactive processing service, drawing board, annotation function, color palette, geometric transformation, data subscription, data cache, data processing, protocol management, analog pixel engine, logic pixel engine, configuration saving, task loading and other functions. The platform interface layer proposes a user-oriented software development method. It is a conduit between the whole system and the outside world. The primary interfaces are Drawing, data, control, and recording interfaces.

The functional application layer realizes a higher level of human-machine dialogue. It includes screen editing, artboard drawing, meta-loading, data modeling, screen organization and other functions.

The primitive component library extends the software at the component level. Perform maintenance work as required. In terms of types, it can be divided into quasi-object elements, logical elements, brush, graph, etc.

**2.2. "Plug-in" function in the spatial model.** In space operation, multiple components' states are often interrelated. A common approach is establishing a "number" or "status" parameter. The central computer data processing service generates the virtual parameter data. Commercial software orders virtual parameters for status presentations are needed. This approach applies to new equipment, models, tasks, and systems. Creating virtual parameters is not suitable for system expansion and state maintenance. The system design of the component connection mechanism only needs to set the correlation of components when the web page is set up. Setting the mode of operation allows the relevant state process to be performed. It is also extended to analog graphic elements and logical graphic elements. The consistency of the system is ensured at the data level. The analogs and logical primitives in the primitive's widget library are inherited from the QtUiLoader interface. In the case of no secondary development, components and layouts can be directly loaded into the business software to achieve the same interface of the business software. The construction of the primitive component is shown in Figure 2.2.

The primitives are inherited from the QtUiLoader. It can be mounted on the frame. The interface of receiving data, subscribing status and disclosing status is realized in this module. Use callback methods to complete data and state processing in the current part.

**2.3. Flow chart of the visual interactive modeling system.** This system is used to design the visual model of the network for different working types and models. They are modeling based on new model tasks and modeling based on historical tasks. Existing working model resources of the same type can be built on top of a historical working model. Only incremental changes are required [8]. The process of the visual topology model is shown in Figure 2.3.

First, the system tasks are classified and compared with the historical tasks in the system. If the match is successful, the historical task is directly loaded. It includes: setting page layout files, task history parameter data, task history theory data, etc. You can make incremental changes and add pages based on historical work. This would significantly reduce the preparation for such missions [9]. If the match is unsuccessful, new task information needs to be entered. The main contents include task name, task code, task parameters, task theory data, task menu structure, etc. After the dialogue between two people, arrange the webpage requirements and set up a new webpage according to the matching results and work requirements.

Secondly, the graphics meta component is initialized. Update the original component list and present the available information to the user. Create a new empty page. Parse the element requirements on the task page. Select the original component and drag it to the empty page [10]. Configure the graphic element style and structure layout. Repeat the previous steps until all graphic elements on the page are filled. Connect

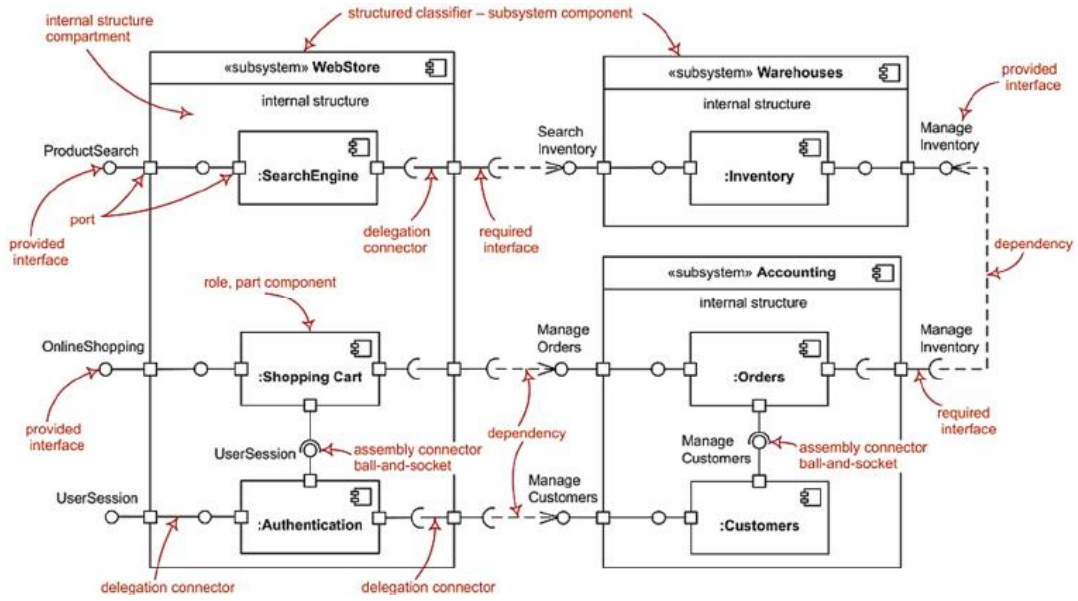


Fig. 2.2: Structure diagram of the element component.

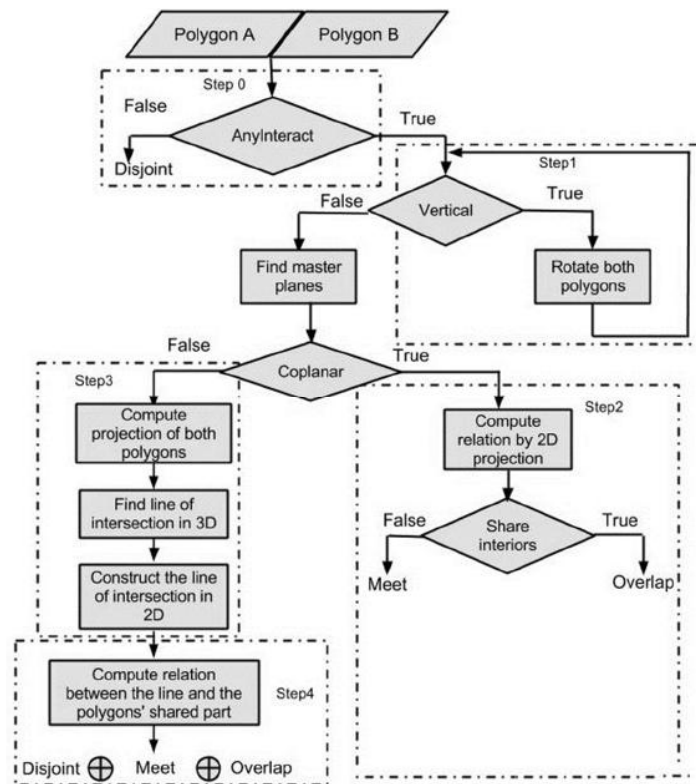


Fig. 2.3: Flowchart of visual topology modeling.

Table 3.1: Efficiency comparison of modeling methods.

	Task preparation time /min	a single page time /min
Previous generation modeling platform	6823	13.65
Visual application topology modeling platform	1344	2.69

the corresponding graph nodes—Configure graphic element associations. Set up logical rules and simulate the behavior effects of objects according to the categories of graphic elements.

Finally, the graphics metadata source is configured, and whether the configuration is correct will be automatically judged. Save the web Settings to the local file when the settings are finished. It is used to load commercial software.

**3. System implementation.** According to the concept of "platform + plug-in," a visualization system of a 3D network model is developed. A series of graphic component libraries are designed for the control platform.

**3.1. Build a primary model development platform.** Firstly, the basic abstracting and extracting business and interface functions are completed from the architecture. The system is based on the Qt designer source code. It is a successor to the Qt localized visual operation interface. Provide user interface file management, editing, page layout, component management, property management and other essential functions. Implement filtering, caching, data subscription, task loading, exceptional attribute management and other functions [11]. The loading framework of logical elements and simulative elements is given. It can carry new symbols and added work management functions—the ability to load a working tree architecture. Build a primary model development platform. Establishing a basic modeling platform cannot realize all the system's functions but only provides the program object-oriented component library and primary data loading platform. In this process, you must enter the task information, real-time data resource configuration information, and historical data resource information in advance. A task configurator must be used to perform related functions—button wake-up function set on the platform.

**3.2. Task Resource Configuration Tool.** A resource setting tool that must perform work to set primary data. This tool can handle the database directly. It maintains and manages the database. The task resource configuration tool can manage all the work. Includes basic Settings for adding, viewing, modifying, and more. The job contains the job name, code, and description information. Manage all the real-time analysis data in the system [12]. Convert live data to subscription data. Provide ordering information for the component, including real-time and historical data. A parallel editing algorithm based on multiple nodes is proposed. This tool can update configuration information to a network database. Online database information can be brought to the local area.

**3.3. Creation of graphics meta component.** This paper completed the system design after the basic model construction and task resource allocation. You need to create a database of graphics metal parts for this article. The graph elements are divided into two types, namely logical graph elements and quasi-object graph elements. The logical graph elements include a decision tree, flow diagram, relationship diagram, dynamic pipeline diagram, equipment diagram and process diagram [13]. The simulated object elements include 14 elements, such as a switch, button, status light, pipeline, line, network equipment, telemetry equipment, external measurement equipment, code table, storage tank, curve chart, line chart, bar chart, pie chart, etc. One of the logical primitives represents the topology composition relationship. Various logical relationships are realized by configuring the elements of logical primitives according to certain logical relationships. The pseudo-object element reflects the connection between the pipe and the device. Lines connect and configure simulated elements through pipes—incremental updates to component libraries based on task requirements [14]. The relevant statistics listed in Table 3.1 are based on a one-person operation with 500 web pages per task. The modeling platform developed based on the visualization application topology modeling method significantly improved the preparation time of the single task and single page configuration. The operator is becoming more and more proficient. Its application effect will also be better.

#### 4. Space data processing algorithm.

**4.1. Sparse point cloud reconstruction.** The SFM algorithm is essentially the projection matrix of the camera. The algorithm uses the relevant positioning information of the camera. 3D information about objects projected backward. After repeated matching, the 3D point cloud of the object and the attitude of the camera are obtained. The projection of  $R$  the three images is  $r_1, r_2, r_2^S G r_1 = 0$ . The geometric limits of the poles are satisfied. The calculation is carried out with  $G$  matrix. Finally, the matrix of the internal and external parameters of the camera is obtained [15]. Combined with the knowledge of triangulation, it can reverse the three-dimensional coordinates of point  $R$ . After obtaining the original projection matrix, the paper continues to add a new projection matrix method for repeated operations. The 3D coordinates and projection matrix are obtained by minimizing the equation of (4.1):

$$\min_{C_i, S_i, X_j} \sum_{i=1}^2 \sum_j ||r_{ij} - A[C_i | S_i] R_j||^2 \quad (4.1)$$

Two feature points of three-dimensional point  $R_j$  of  $(r_{1j}, r_{2j})$  are matched.  $A$  is the matrix of the camera's internal parameters. Where  $[C_i, S_i]$  is the external parameter of the camera.

First, two cameras with the most feature points in the image are selected as targets; The five-point method (RANSAC) was used to identify the camera parameters. Two cameras observe feature points. The spatial coordinates of the feature points are obtained by trigonometry [16]. The beam current adjustment method is used to optimize it. New cameras are added in turn, and the feature points observed by them are used as the target optimal solution. After adding the above parameters, the overall beam adjustment is completed. This method is repeated until other cameras find no new feature points. Some optimizations have been made for speed and robustness:

- 1) The singularity with high projection error is removed each time the optimization is performed—one more adjustment until there are no more singularities.
- 2) Initialize multiple cameras simultaneously. First, add a camera with  $\lambda$  maximum corresponding points. Add at least one feature point that corresponds to  $0.75\lambda$ .

**4.2. Poisson reconstruction.** Poisson reconstruction is based on the fact that the equivalent function of the indicator surface reconstructs the points in the point cloud model. It is an implicit function method. This paper transforms the surface reconstruction problem of directed point sets into a spatial Poisson problem. Set the point set in the point cloud as a sample set  $l \in L$ . Each sample contains a point and an inward average vector. Assume the set of points is on the model's surface [17]. An approximate representation of the model is obtained by estimating the indicator function of the model. The seamless triangular approximation to the model surface is realized by reconstructing the contour surface. The reconstruction program is divided into:

Step 1 is the definition of the octree. Octree is a kind of data structure used to represent the three-dimensional information of objects in space [18]. An octree structure based on a point set  $L$  is proposed. The maximum value of an octree is  $Z$ . Every node in an octree is  $P$ . Set the function  $G_P$  to represent the function for each node.  $P.B$  stands for the middle node  $P$ .  $P.E$  represents the width of a node  $P$ .  $v$  stands for a subject. The extension of  $G_P$  is

$$G_P(v) = G \left( \frac{v - P.B}{P.E} \right) \frac{1}{P.E^3} \quad (4.2)$$

In order to make  $U$  can be expressed by the linear sum of node functions, the basis function of  $G$  is filtered by  $n$  times envelope. When the  $C$  value increases, the properties are similar to Gaussian filters.

Step 2 is the calculation of the vector field. The approximation of the functional gradient magnetic field of this index of the child node is calculated by the formula (4.3):

$$U(v) = \sum_{l \in L} \sum_{P \in N_g \text{ br } Z(l)} \varepsilon_{P,L} G_P(v) l.N \quad (4.3)$$

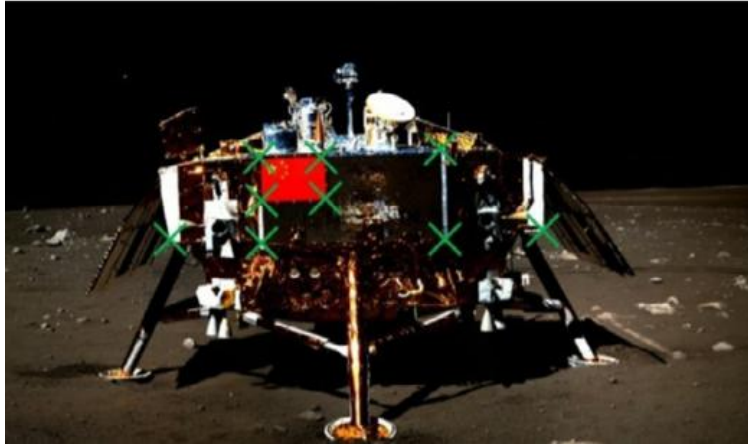


Fig. 5.1: Feature points extracted from the image.

$N_{gbr}Z(l)$  represents the 8 nodes closest to the sampling point  $l.r$  and  $Z$ . Where  $\varepsilon_{P,L}$  is the weight of third-order linear interpolation.  $l.N$  represents an average vector. The vector field  $U$  is generated by this method.

Step 3 is the solution to Poisson's equation. Use the vector field  $U$  obtained in (3). The index function  $\rho$  of the surface can be obtained by the inverse solution of Poisson's equation. The problem is solved in detail by using the Laplacian iterative matrix.

Step 4 is the construction of the equivalent height surface. The step of surface reconstruction is essentially to extract the surface [19]. The position of the sampling point must be estimated and averaged to obtain the surface  $\partial M$ . The equivalent surface is derived by formula (4.4) and formula (4.5):

$$\partial M = \{v \in C^3 | \rho(v) = c\} \quad (4.4)$$

$$c = \frac{1}{|L|} \sum_{l \in L} (l.r) \quad (4.5)$$

$|L|$  is how many points there are in a set of points.

## 5. Experimental design.

**5.1. Experimental Data.** The measured data test the performance of the software. C++ writes the visualization program of 3D reconstruction. The operating system is Windows 2010. The processor is an Intel i5—speed 7200r/min. In the image acquisition process, this paper simulates the Angle of image acquisition between satellites. Using mobile phones to get images of the lunar probe [20]. In this paper, the image data of the lunar rover are obtained continuously at a certain Angle, and the image data is input into the software platform.

**5.2. Result Analysis.** Use a ground-based application to transmit high - and low-density reconfiguration commands to a cloud application on a satellite. The resulting 3D reconstruction file is stored in the PLY format. The 3D reconstruction results are displayed in mobile phones and PCS applications. The results of sparse, dense, and Poisson reconstruction of point clouds are demonstrated by MeshLab [21]. Figure 5.1 shows extracting feature point information from non-collaborative objects in 3D reconstruction processing.

The effects of different numbers of experimental images on the reconstruction results were studied. The paper selected 10, 15, and 20 images as the result of the 3D reconstruction (Figure 5.2). When the number of images is 10, the result of 3D reconstruction is relatively poor. The results of 3D reconstruction were better in the case of 15 and 20 photos. And there are more and more photos, and they take longer and longer. The

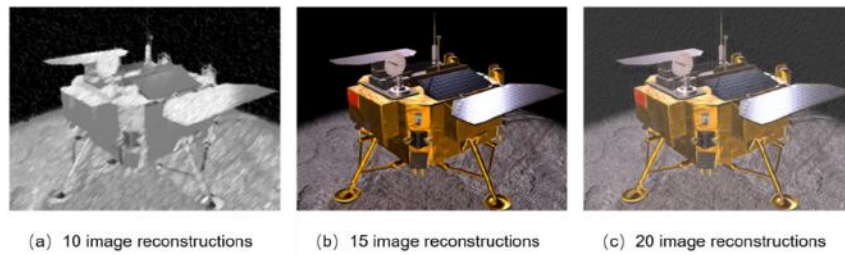


Fig. 5.2: Reconstruction results of dense point cloud.



Fig. 5.3: Poisson reconstruction results.

ten photos were restored for a total of about 17 minutes. When the number of images is 15, the reconstruction time is about 30 minutes. When the number of images is 20, the reconstruction time is about 40 minutes. It was reconstructed in 3D with 15 images [22]. It can also be seen from Figure 5.2 that dense point cloud reconstruction has the disadvantages of poor visualization and sparse generated point cloud. To get better display results, intensive reconstruction needs to be further improved.

The paper selected 15 images for 3D reconstruction. The Poisson reconstruction method is used for further optimization based on the reconstruction effect shown in Figure 5.2. The defects in the reconstruction process of the original dense point cloud can be effectively remedied by grid processing. Form a unified and compact surface. The results of the refactoring are shown in Figure 5.3. Through comparative analysis, it can be found that the software system completely restores the structure of the lunar probe. This means that the reconstruction is working very well.



**6. Conclusion.** This paper mainly studies the visual topology modeling method of space missions and analyzes and compares the two feasible methods of the current technical route. It provides a new, extensible framework for developing the spatial visualization topology model. This paper introduces a space command display software development scheme on the QtUiLoader platform. The development of the system is easy, and the development period is short. This algorithm is used in space command display software to complete the visualization of the space control interface and the creation of a topological structure diagram. The system completes the feature point extraction, matching and camera position calculation of non-cooperative object images based on the satellite-borne optical camera. A solution model is obtained based on a sparse point array and camera coordinates. It can solve the problem of vector relationship between satellite and non-cooperative object. The system can achieve accurate 3D reconstruction of non-cooperative objects, which meets the requirements of human eye observation and capture. This project provides the theoretical basis and technical support for developing the spatial visualization topology model. However, further work still needs to be carried out to realize human-computer interaction in the system and organically combine the visual topology model with the commercial system simulation.

**Acknowledgments.** Guangdong Provincial Science and Technology Innovation Strategy Special Fund Project (pdjh2023b0984); Scientific Research Project of the Science and Technology Department of Jingchu Institute of Technology(YB202319). General Science and Technology Plan Project of Jingmen Science and Technology Bureau(2023YFYB085).

#### REFERENCES

- [1] Gui, G., Zhou, Z., Wang, J., Liu, F., & Sun, J. (2020). Machine learning aided air traffic flow analysis based on aviation big data. *IEEE Transactions on Vehicular Technology*, 69(5), 4817-4826.
- [2] Qiu, R., Hou, S., Chen, X., & Meng, Z. (2021). Green aviation industry sustainable development towards an integrated support system. *Business Strategy and the Environment*, 30(5), 2441-2452.
- [3] Meister, P., Miller, J., Wang, K., Dorneich, M. C., Winer, E., Brown, L. J., & Whitehurst, G. (2022). Designing three-dimensional augmented reality weather visualizations to enhance general aviation weather education. *IEEE Transactions on Professional Communication*, 65(2), 321-336.
- [4] Ayala, R., Ayala Hernández, D., Sellés Vidal, L., & Ruiz Cortés, D. (2021). openSkies-Integration of Aviation Data into the R Ecosystem. *The R Journal*, 13 (2), 590-599.
- [5] Mao, R., Li, G., Hildre, H. P., & Zhang, H. (2021). A survey of eye tracking in automobile and aviation studies: Implications for eye-tracking studies in marine operations. *IEEE Transactions on Human-Machine Systems*, 51(2), 87-98.
- [6] Li, S., Zheng, P., & Zheng, L. (2020). An AR-assisted deep learning-based approach for automatic inspection of aviation connectors. *IEEE Transactions on Industrial Informatics*, 17(3), 1721-1731.
- [7] Puranik, T. G., & Mavris, D. N. (2020). Identification of instantaneous anomalies in general aviation operations using energy metrics. *Journal of Aerospace Information Systems*, 17(1), 51-65.
- [8] Kanavos, A., Kounellis, F., Iliadis, L., & Makris, C. (2021). Deep learning models for forecasting aviation demand time series. *Neural Computing and Applications*, 33(23), 16329-16343.
- [9] Memarzadeh, M., Matthews, B., & Templin, T. (2022). Multiclass Anomaly Detection in Flight Data Using Semi-Supervised Explainable Deep Learning Model. *Journal of Aerospace Information Systems*, 19(2), 83-97.
- [10] Sun, J., Gui, G., Sari, H., Gacanin, H., & Adachi, F. (2020). Aviation data lake: Using side information to enhance future air-ground vehicle networks. *IEEE Vehicular Technology Magazine*, 16(1), 40-48.
- [11] Elakramine, F., Jaradat, R., Hossain, N. U. I., Banghart, M., Kerr, C., & El Amrani, S. (2021). Applying systems modeling language in an aviation maintenance system. *IEEE Transactions on Engineering Management*, 69(6), 4006-4018.
- [12] Shome, S., & Verma, S. (2020). Financial distress in Indian aviation industry: Investigation using bankruptcy prediction models. *Eurasian Journal of Business and Economics*, 13(25), 91-109.
- [13] Kim, G., & Sterkenburg, R. (2021). Investigating the effects aviation fluids have on the flatwise compressive strength of Nomex® honeycomb core material. *Journal of Sandwich Structures & Materials*, 23(1), 365-382.
- [14] WANG, S., CHEN, H., DING, L., SUI, H., & DING, J. (2023). GAN-SR Anomaly Detection Model Based on Imbalanced Data. *IEICE TRANSACTIONS on Information and Systems*, 106(7), 1209-1218.
- [15] Puranik, T., Harrison, E., Chakraborty, I., & Mavris, D. (2020). Aircraft performance model calibration and validation for general aviation safety analysis. *Journal of Aircraft*, 57(4), 678-688.
- [16] Babae, M., Gheidar-Kheljani, J., Khazae, M., & Karbasian, M. (2021). Prediction of Failure Time and Remaining Useful Life in Aviation Systems: Predictors, models, and challenges. *International Journal of Reliability, Risk and Safety: Theory and Application*, 4(2), 97-106.
- [17] Avci, C., Tekinerdogan, B., & Athanasiadis, I. N. (2020). Software architectures for big data: a systematic literature review. *Big Data Analytics*, 5(1), 1-53.
- [18] Walia, S., Sharma, D., & Mathur, A. (2021). The impact of service quality on passenger satisfaction and loyalty in the Indian aviation industry. *International Journal of Hospitality and Tourism Systems*, 14(2), 136-143.

- [19] Jiang, Y., Niu, S., Zhang, K., Chen, B., Xu, C., Liu, D., & Song, H. (2021). Spatial-temporal graph data mining for iot-enabled air mobility prediction. *IEEE Internet of Things Journal*, 9(12), 9232-9240.
- [20] Gogalniceanu, P., Calder, F., Callaghan, C., Sevdalis, N., & Mamode, N. (2021). Surgeons are not pilots: is the aviation safety paradigm relevant to modern surgical practice?. *Journal of Surgical Education*, 78(5), 1393-1399.
- [21] AKPINAR, M. T., HIZIROĞLU, K., & CODAL, K. S. (2023). From Descriptive to Prescriptive Analytics: Turkish Airlines Case Study. *Adam Academy Journal of Social Sciences*, 13(1), 99-125.
- [22] Lee, B., Brown, D., Lee, B., Hurter, C., Drucker, S., & Dwyer, T. (2020). Data visceralization: Enabling deeper understanding of data using virtual reality. *IEEE Transactions on Visualization and Computer Graphics*, 27(2), 1095-1105.

*Edited by:* Zhigao Zheng

*Special Issue on:* Graph Powered Big Aerospace Data Processing

*Received:* Jul 18, 2023

*Accepted:* Aug 25, 2023

---

## 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.