Scalable Computing:
Practice and Experience

Scientific International Journal
for Parallel and Distributed Computing

ISSN: 1895-1767

Universitatea de Vest din Timişoara

Volume 24(2) June 2023

SUBSCRIPTION INFORMATION: please visit `http://www.scpe.org`

Scalable Computing: Practice and Experience

Volume 24, Number 2, June 2023

TABLE OF CONTENTS

Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2023 SCPE. Volume 24, Issues 2, pp. 81–95, DOI 10.12694/scpe.v24i2.2078

TECHNOLOGY ENABLED INTELLIGENT SOLUTION IN HUMAN RESOURCE MANAGEMENT FOR SMART CITIES

GARIMA VIJH *, RICHA SHARMA †, AND SWATI AGRAWAL ‡

Abstract. The foundation of smart cities is based on an autonomous and decentralized architecture, which consists of sophisticated information and communication technologies (ICT) in convergence with technology enabled solution to improve the business management process in industry 4.0. This study tends to examine the adoption of blockchain technologies (DLT) in the human resource management (HRM) of organizations in building solutions for IOT (Internet of things) smart cities.

The current study explores a unique set of factors selected from the extensive literature and acquired information from fifteen experts having significant experience of blockchain technology in their respective organizations. An integrated fuzzy analytic hierarchy process (F-AHP) is applied to prioritize the identified success factors. Further, the modified decision-making trial and evaluation laboratory (M-DEMATEL) method is utilized to represent the complicated causal relationships among different sub-factors on blockchain-HRM integration.

The findings show the application of blockchain will foster a paradigm change in IOT based smart communities, where recruiters verify the candidate credentials including education, skills, and work experience. The payroll managers would determine the more effective way to make work less complex and moderate, enabling timelier payments to global employees. Furthermore, DLT would enhance the employee learning records and update the real-time information in HRM database technologies. Thus, providing a detailed guide for future Industry 4.0 developers about how blockchain can improve the next generation of industrial applications.

The developed method can help the decision-makers and provide a foundational view to examine the benefits of implementing blockchain technology in the HRM setting of an organization before they choose to integrate in order to enhance Industry 4.0 technologies.

This research will be a novel attempt to synthesize the key factors and subfactors about technology enabled solution within the intelligent HRM process, shedding light to rethink HRM strategies to incorporate blockchain technology in organizations.

Key words: Blockchain Technology (BCT), Human Resource Management (HRM), Distributed Ledger Technology (DLT), Fuzzy AHP (F-AHP), Modified DEMATEL (M-DEMAEL)

1. Introduction.

With the ubiquitous implementation of the Internet of things (IOT) and numerous Industry 4.0-based applications such as cloud computing (CC), the Internet of services (IOS), cyber-physical systems (CPS), and data analysis techniques have created an interconnected world in which sensors and actuators perceive, compute, and transfer data to automate the industrial network in smart city. These advanced technologies provide countless intelligent solutions for boosting automation, efficiency, security, precision, and scalability in a variety of industrial and business segments. Blockchain technology is considered an extensive innovative synergist revolutionizing the structure of interaction, creation or valuation between businesses, customers, societies, partners, and individuals [1]. According to the world economic forum predictive analysis, 10% of the global GDP would be on hybrid blockchain technologies by 2025. Gartner [2] determines that blockchain monetary worth will exceed $3.1tn by 2030. There are various companies like Google, Amazon, Facebook, and IBM that are leveraging the blockchain in their core processes to deliver valuable solutions to their clients and employees [3, 4] .

Blockchain is an emerging technology that adds significant value in the HRM domain and streamlines various activities of HR professionals. Human resource management is the most integral component of the organizations entangled with coordinating substantial tasks varying from recruiting and retaining the best talent, processing payroll, regulatory compliance reporting, providing a safe working environment, activating training & development programs, and offering enormous incentives to achieve organizational objectives [5]. Blockchain-based applications will proliferate the human resource industry with immense benefits, that will

*Amity University, Noida, Uttar Pradesh, India (garimavijh672@gmail.com).
†Amity University, Noida, Uttar Pradesh, India (rsharma25@amity.edu)
‡Bennett University, Greater Noida, Uttar Pradesh, India (swati.02@gmail.com)

allow organizations and firms to develop and host HR-related activities on the blockchain. Blockchain networks can serve as an imperative instrument for the verification of individual identities, educational background checks, and employment histories, thus leading to attract quality talent for the job position [6]. Additionally, BCT has the potential to provide real-time verification of cross-border payments, without the need for intermediaries such as banks, third parties and clearinghouses, thus workers can be paid instantly with the use of smart contracts [7].There is also a possibility that the performance appraisal model associated with rewards, recognition and salary increments can be re-designed with the distributed ledger platform in an organization [8].Thus, shifting the conventional role to a more strategic role leading the HRM core functions with more realistic and valuable decision support information.

The functioning of HRM in organizations utilising Industry 4.0-supported technical initiatives creates the research gap in the existing system. Thus, HRM system optimization, facilitated by cutting-edge technologies like blockchain promoting the flexibility and agility for improving HRM core activities [9]. Consequently, this research contributes to the organization's HRM domain by addressing the following questions:

RQ1: What are the prominent determinant and sub-determinant for blockchain implementation to enhance the human resource management functions of an organization?

RQ2: How are the stated factors prioritized for the integration of blockchain with HRM practices?

RQ3: How are the selected determinants and sub-determinants interrelated for improving HRM practices through the usage of blockchain?

To resolve abovementioned research questions, this paper attempts to implement an integrated Fuzzy Analytic Hierarchical process and modified DEMATEL methodology to determine critical success factors and sub-factors to uncover their applicability. The objective is to establish an interpretive structural model representing significant success factors while simulating technology based solution using fuzzy approach. Thus, this paper contributes to explore the opportunities regarding the deployment and usage of blockchain technology, particularly for the HRM sector in the smart city to build sustainable HRM ecosystem(services).

2. Literature Review.

2.1. Theoretical foundation. The Information system (IS) adoption literature has been developed by several researchers such as the Innovation Diffusion Theory (IDT), and Technology Acceptance Model (TAM), which have been instrumental in describing the adoption of a variety of IS in their corresponding organizational work [10]. According to the recent challenges faced by HR managers/consultants, the core construct of the unified theory of acceptance and use of Technology (UTAUT) includes behavioural intention to use the technology in the organizational setting [11]. The main determinant of the model is performance expectancy, effort expectancy, social influences and promotion conditions in predicting the user's intention toward blockchain implementation [12]. UTAUT was further modified to include more contextual factors, such as price value, habit, and hedonic motivation thus forming UTAUT2 [13]. The expansion of UTAUT2 provides a foundation for precisely analyzing the adoption and use of emerging blockchain technologies in HRM [14, 15]. This theoretical framework serves as guidance (underpinning theory) for the evolution of research propositions in blockchain technology to be applied in human resource management functions. Based upon the UTAUT/UTAUT2 description, the four building blocks are deployed in the HR department: talent acquisition, payroll processing, regulatory & Compliance management, and performance management, thus giving insights about recreating HRM services and enhancing the path of growth, improvement, and efficiency.

2.2. Blockchain application and its characteristics. In this section, we briefly describe the related work on the application of BCT in various industries and then summarize the possible advantages and disadvantages of BCT. Lastly, we explain in detail the utilization of BCT in Human Resource Management, followed by the identification of success factors.

2.2.1. Blockchain overview. Blockchain technology is one of the latest innovations across industries, practitioners, and academia. The subject accelerated enriching absorption post-2016 after the acceptance of cryptocurrencies as a non-fiat decentralized digital currency and payment infrastructure system that operate at a global level to prevent forgeries, false disputes, double spending etc [16]. Blockchain technology has versatile uses apart from recording financial transactions. This technology can be utilized for asset tracking and registration, decentralized voting and governance, healthcare information, transportation, digital supply

chain and much more [17, 18]. Blockchain is explained as a decentralized, transactional database that stores a congruent, immutable, and chronological log of transactions between network participants [19, 20]. The record of validated information gets sealed by a cryptographic algorithm, linking blocks of data in form of a chain to ensure data integrity and standardized arrangement for the data access [21]. Each block contains the encrypted data and reference links to the preceding block hence maximizing the data preservation with a coded security system. Furthermore, all participants hold a copy of the ledger that cannot be altered retroactively, thus providing transparency, trustable and tamper-proof records [22]. Thus, the extensive literature show that the integration of blockchain technology with HRM practises could provide tremendous possibilities to deliver HR solutions at low expense associated with organization and development of sustainable smart cities. Table 2.1 and 2.2 highlights the benefits or essential blockchain technology enablers (BCTE's) that motivate organizations to incorporate blockchain in their HRM practices. However, there are a few challenges the organization initially face while the successful implementation of blockchain technology.

2.2.2. Impact of blockchain technology in HRM. Blockchain is cultivating momentous excitement in many aspects of industries by virtue of its design and architecture. Sivathanu [39]demonstrated that blockchain technology has great potential to upend HR functions that are not just restricted to talent management such as recruiting, developing and retaining new-age talent for achieving high performance at an organizational level. Blockchain will deeply impact HRM department roles and responsibilities, thus going to disrupt big picture areas like hiring, and payroll with its impressive services [40].The hiring process has become a daunting task for the human resource department of the organizations due to the false information provided by the job applicants on their CVs. Nearly,75% of hiring managers identified wrong credentials on the resume of applicants while applying for a suitable job [41]. Some job seekers hoax their academic degrees, employment letters, job titles and deliberately exaggerate their skills. The prime concern of the current organization is tracking down the potential candidate that could match the requirements of the vacant positions by witnessing their education, expertise and achievement [42].

The disruptive blockchain technology reduces the occurrence of unscrupulous entries and employers before hiring a candidate can get their most precise delineation and contestant's credentials. Thus, the integration of blockchain will verify the data submitted by the employee will provide genuine, transparent information, reduce costs and make it more effective [43]. Upon successfully verifying the candidate data, organizations can quickly release the offer letter, thereby saving significant time, effort and cost in the recruitment process. Concurrently, individuals can enhance their employability by voluntarily sharing their data with the recruiters as it improves employment opportunities and creates a culture of trust between the employee and the employers [44]. Moreover, recruiters can access the information about each candidate in real-time rather than receiving traditional resumes or viewing career networking websites like LinkedIn in near future. Thus, transforming the HRM function with a blockchain new intelligent hiring game plan.

Blockchain technology also emerged as a potential disruptor in the payment solutions arena. The greatest challenges faced by the HRM department in maintaining high volume financial operations. The procedure that employees value the most such as payroll process, taxes, cross-border transactions, and managing regulatory compliance can be made faster, secure and transparent with the advent of blockchain in the gig economy [45]. The application of blockchain in payments system ensures that participating parties can share sensitive information in a verified manner without relying on central authorities or third-party intermediaries, thus eliminating human errors and developing a trusted network [46].To sum up, HR managers would not be required to run a monthly payment or contact their company's bank, thus ensuring quick and low transmission costs for its customers or regulators. Instead, open real-time blockchain ledgers will help HRM head with invoice tracking as well as transaction distribution, invoicing, and reporting.

The next big space where blockchain can be used to automate the performance evaluation system. The performance management cycle is crucial for all employees in an organization because it is associated with the recognition, salary hike, training and career growth. With the introduction of online training courses, the availability of digital badges, and the vast array of courses, education, and academic achievements, HRM teams are finding it more difficult to maintain track of each employee's day-to-day activities in the organization [47]. Therefore, HR professionals are looking for new solutions to keep employees engaged and productive. The application can be developed to create a decentralized performance appraisal measurement system using

Table 2.1: Summary of benefits and limitations of blockchain adoption

Sr. No	Benefits/ BCT Enabler's	Description	Refs.	Challenges	Description	Refs.
1	Decentralized data records	In a decentralized network, data is not controlled by any large, centralized server but rather evenly distributed across different nodes (Peer to Peer Network).	[23]	Low Scalability	The blockchain scalability issue is connected to the fact that the size and frequency of records (or blocks) in the blockchain are restricted. The blockchain's blocks continue to expand in size as it is used, and each transaction takes longer to execute.	[32]
2	Consensus-based and Trustworthy Insights	All transactions are validated to concerned participants in real-time, thus ensuring the whole system is fault-tolerance.	[24]	Security enhancement	A 51% attack on the blockchain through node hacking may occur as a result of the lack of validator nodes. The entire data is vulnerable if one of the participant's private keys is lost or stolen. Hence, the system has no safety mechanism to provide additional security.	[33]
3	Immutability	The cryptographic hash algorithm ensures transaction data stored cannot be altered, corrupted, or removed thus, helps in maintaining a high level of robustness and trust.	[25]	High Energy Consumption	The consumption of power in the Blockchain is comparatively high due to mining activities. Every time the ledger is updated with a new transaction, the miners need to solve the problems which means excessive energy costs.	[34]
4	Transparency	Every user has identical copies of the ledger, to monitor and analyze the state of a transaction in its lifecycle, therefore, eliminating the information gap.	[26]	Regulatory and Legal Formalities	Due to the lack of laws and regulations governing blockchain and smart contracts, many organisations may not implement blockchain-based solutions.	[35]
5	Removal of non-value adding intermediary or cost reduction	BCT eliminate intermediaries or third parties hence maximizing transparency, cutting down the overall processing cost and offering a faster settlement of transactions.	[27]	Privacy Leakage	Blockchain is an open ledger that is accessible to all users to increase transparency and eliminate trust issues. But it becomes a liability if applied in a sensitive environment.	[36]

blockchain, smart contracts and crypto coins, thus providing a rewarding experience for an overall contribution of employees towards the organization over a given interval of time [48].

In the conventional centralized HRM transaction model, each transaction needs to be validated through

Table 2.2: Summary of benefits and limitations of blockchain adoption

Sr. No	Benefits/ BCT Enabler's	Description	Refs.	Challenges	Description	Refs.
6	Agile Information Sharing	The issues that hamper the HR process such as delays in payments, conflicts between parties, fraudulent practices and data vulnerability can be minimized with the help of blockchain. Thus, BCT facilitates sharing, exchanging and integration of information across all the network users.	[28]	Interoperability	Most blockchains operate on their own and do not interact with other peer networks as they cannot send and receive information from another blockchain-based system. It is one of the core reasons organizations are not adopting blockchain technology.	[37]
7	Secured database	The identities of users are kept anonymous throughout transactions to protect the security of data. The various functions such as deletion, updating etc. cannot be performed on electronic records, thus prevents from fraudulent actions.	[29]	Lack of Expertise Knowledge	Even though blockchain developers are in high demand, an extreme shortage of blockchain specialists and developers is a major concern for all organizations. The absence of trained and knowledgeable developers for managing and solving the complexity of peer-to-peer networks contributes to the slow rate of development.	[38]
8	Secured database	The identities of users are kept anonymous throughout transactions to protect the security of data. The various functions such as deletion, updating etc. cannot be performed on electronic records, thus prevents from fraudulent actions.	[30]			
9	Simplifies Audibility	It helps to control illicit activities, streamline audit processes, and improve financial reporting.	[31]			

the central trusted agency (e.g., the central bank or intermediaries), inevitably resulting in cost and performance bottlenecks at the central servers. However, a decentralized blockchain-based HRM system ensures a protection mechanism for sensitive information, with an automated verification credential process, building an interplanetary file system to track and store transactions safely without the involvement of third parties, thereby diminishing the usage of the traditional architecture approach [49] [50]. Thus, becomes crucial for the practitioners in the HRM to identify the drivers of the adoption of BCT and comprehend the co-relation among each other. Based on the above literature, a consolidated list of significant factors is framed and divided into Talent Acquisition, Payroll processing, Regulatory compliance procedures and Performance for the

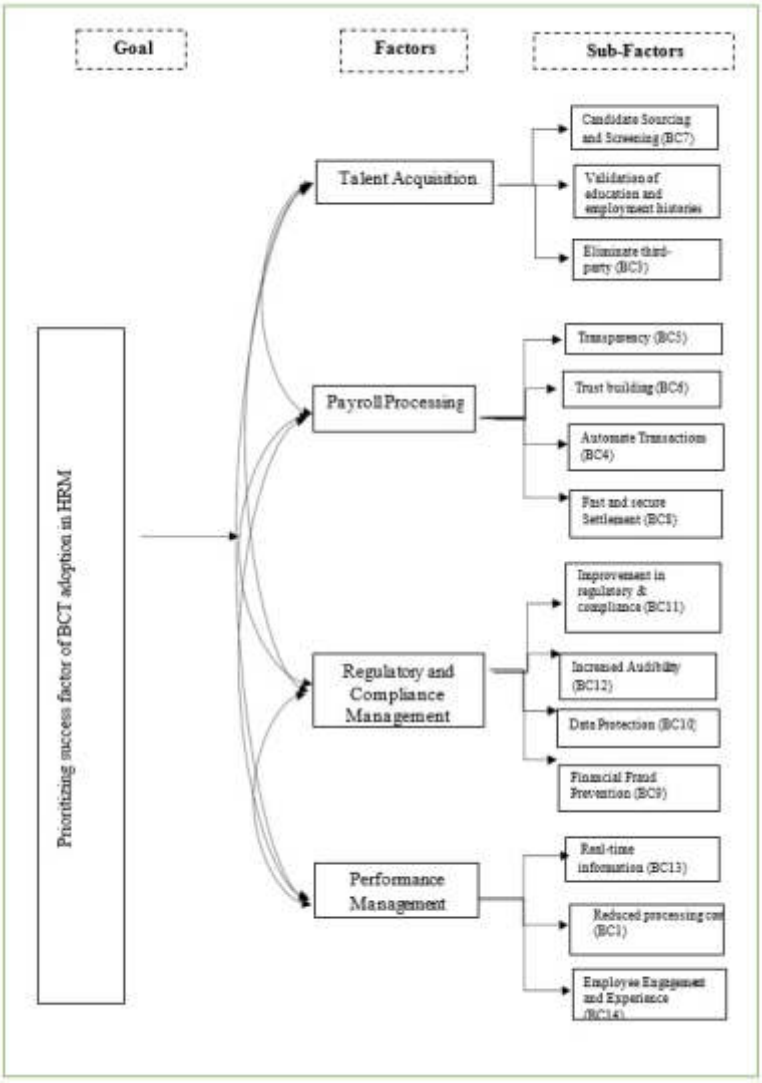


Fig. 2.1: AHP model for Blockchain Technology adoption in Human Resource Management

proposed model. This involves designing hierarchical levels and links between the chosen BCT enablers using a combination of the Fuzzy Analytic Hierarchy Process (F-AHP) and the Decision-Making Trial and Evaluation Laboratory (DEMATEL) methodology. Therefore, an attempt is being made to use various techniques and strategies that will encourage, engage, and create a high-performing management system that endorses the usage of blockchain technologies in the HRM sector at the workplace as depicted in Figure 2.1.

3. Proposed Methodology. Multi-criteria decision-making method [51] performs the analysis for determining the best criteria and structure to solve the complex multiple set problem.

The analytical hierarchal process (FAHP) is broadly used in various applications of multi-criteria decision analysis to discover solution of complicated issues [52]. However, AHP method have been impacted by certain flaws, such as judgements based on unbalanced scales, imprecision, uncertainty, and the biases of decision makers. To address this issues, fuzzy theory has been incorporated into the AHP, as supported by the relevant literature.The F-AHP approach arranges the factors and sub-factor according to their priority or rank. In addition, cause–effect correlations between factors are examined using modified-DEMATEL in this study. In

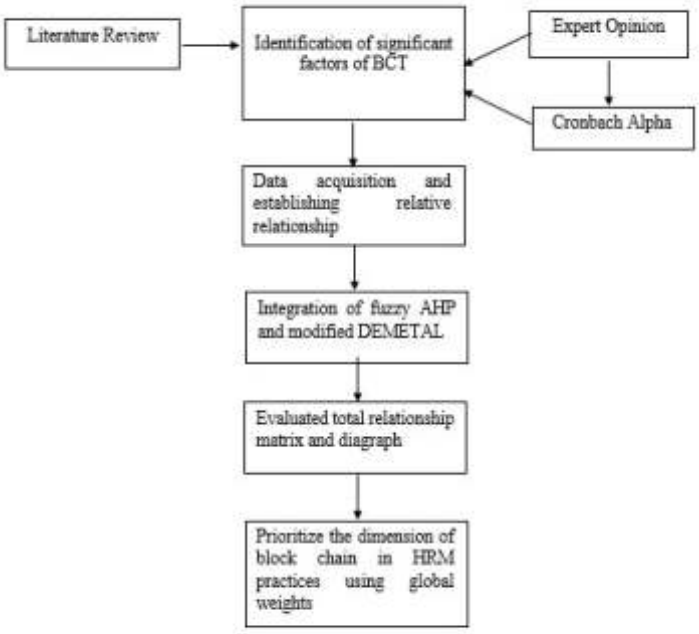


Fig. 3.1: Flow Process of Proposed Methodology

order to avoid uncertainties of human ideas and decisions in a dynamic environment, modified-DEMATEL has been preferred over standard DEMATEL.Therefore,modified-DEMATEL is applied to compute the significant factors and subfactors in order to analyze the inter-dependent relationship[53]. It provides the cause-and-effect mapping of critical significant factors on each other through digraph. The flow of the process has been depicted in Figure 3.1.

3.1. Cronbach Alpha. The critical significant factor has been identified from an extensive literature review and field survey. The survey was developed using a unique set of factors and subfactors from the existing literature and verified by the experts. In this paper, Cronbach Alpha has been used for assessing the strength of consensus agreement among the experts. A total of 15 HRM head/managers or consultants in the organization having deemed knowledge about the acquisition of blockchain in the HRM domain were approached to participate in the survey and were also assured of the confidentiality of their response if they consent to participate. The reliability or inner consistency of the designed questionnaire is determined with the help of Cronbach Alpha [54]. The test revealed a reliability score of 0.83, hence the survey questionnaire was found to be reliable.

3.2. Fuzzy AHP (F-AHP). Cheng et al.[55] presented the F-AHP technique enabling the extent analysis using a triangular fuzzy comparison matrix to determine a crisp priority vector. F-AHP technique is a more methodical approach than the other MCDM approaches to capture a human's perception of ambiguity when complicated multi-criteria decision-making situations are considered for the decision-making process. Due to this ability, F-AHP tends to be advantageous for strategizing and determining judgements to acquire more crisp information for accurate results. However, F-AHP is a more intricate method that requires numerical calculations in determining composite priorities than conventional AHP. In this study, the F-AHP method is employed to obtain the critical factor and subfactor weights to perform a comparison of the different judgement of expert's opinions [56]. Let object set and goal set be represented $Y = \{ y1, y2. \ldots . yn\}$ as and $Z = \{ z1,$

z2 … . . zm} . The 'n' extent analysis values can be acquired using LJ ($j = 1, 2m$) depicting triangular fuzzy numbers.

The detailed step process is discussed as follows:

1. Determine the value of fuzzy synthetic weights with respect to the i_{th} object through Eq. 3.1

$$Q_i = \sum_{j-1}^{m} L_{z_i}^j \otimes \{\sum_{i-1}^{n}\sum_{j-1}^{n} L_{z_i}^j\}^{-1} \tag{3.1}$$

The fuzzy addition operation of 'n' extent analysis value is evaluated for a particular matrix as shown in Eq. 3.2. Further, the inverse of the vector is given by Eq. 3.3.

$$\sum_{j-1}^{m} L_{z_i}^j = \left(\sum_{j-1}^{m} r_j, \sum_{j-1}^{m} h_j, \sum_{j-1}^{m} h_j\right) \tag{3.2}$$

$$\left\{\sum_{i-1}^{n}\sum_{j-1}^{m} L_{z_i}^j\right\}^{-1} = \left(\frac{1}{\sum_{i-1}^{n}\sum_{j-1}^{m} r_j}, \frac{1}{\sum_{i-1}^{n}\sum_{j-1}^{m} h_j}, \frac{1}{\sum_{i-1}^{n}\sum_{j-1}^{m} h_j}\right) \tag{3.3}$$

2. Calculate the degree of possibility (DOG) between $Q1$ ($r1$, $h1$, $v1$) and $Q2$ ($r2$, $h2$, $v2$) fuzzy synthetic extent as defined in Eq. 3.1. The magnitude of likelihood between the two fuzzy synthetic extent is achieved through Eq. 3.4.

$$t(Q2 \geq Q1) = \sup g \geq y \exp \min(\mu q2(g), \mu q1(y))] \tag{3.4}$$

It can also be expressed using Eq. 3.5 and Eq 3.6 respectively.

$$t(Q2 \geq Q1) = hgt(Q1 \cap Q2) = \eta q2(k) \tag{3.5}$$

$$\eta_{q2}(k) = \begin{cases} 1 & \text{if } h2 \geq h1 \\ 0 & \text{if } r1 \geq r2 \\ \frac{r1-v2}{(h2-v2)-(h1-r1)} & \text{otherwise} \end{cases} \tag{3.6}$$

3. Measure the degree of possibility for convex fuzzy number is more than k, q_i ($i = 1, 2, ..., k$) can be depicted through Eq. 3.7.

$$t(q \geq q1, q2, ...qk) = t[(q \geq q1) \text{ and } (q \geq q2)... \text{ and } ...(q \geq qk)] = \min t(q \geq qi), i = 1, 2, 3...k \tag{3.7}$$

4. Determine normalized vector using Eq. 3.8. Assume that

$$b(B_i) = \min t(q \geq qi) NV = (d(B1), d(B2), ...d(Bk))^z \tag{3.8}$$

NV shows the non-fuzzy number computed for each comparison matrix.

3.3. Modified DEMATEL technique. Battelle Memorial Institute Geneva Research Centre proposed DEMETAL techniques for exploring and visualizing the modelling structure that emphasize the interdependent relationship with diagrams and matrices [57]. The detailed associated process of the M-DEMATEL method is described as follows:

1. Compute the average direct relation matrix D from the judgement of experts considering the impact of every factor on other factors. Assume that the study contains a set of factors as D and their mathematical relation e_{ij} $(i, j \in \{1, 2, 3.(m-1), m\}$ determined using pairwise comparison. So, e_{ij} shows the degree by which one factor or subfactors influence the other significant factor.

2. Normalization of direct relation matrix D such that the values can be acquired between 0 and 1 through Eq. 3.9 and 3.10 respectively.

$$O = K \cdot D \tag{3.9}$$

$$O = \min\left(\frac{1}{\max_{1\le i\le m}\sum_i = 1|a_{ij}|} . \frac{1}{\max_{1\le j\le m}\sum_j = 1|a_{ij}|}\right) \tag{3.10}$$

Here, $i, j \in \{1, 2, 3., (m-1), m\}$ where m shows the count of factors in this study.

3. Generate the normalized total direct relation matrix S (here S= identity matrix) depicted through Eq. 3.11.

$$S = M + M2 + M3 + M4 + ... = \sum_{M_{i=1}^{i}}^{\infty} M(1-M) - 1 \tag{3.11}$$

4. Determine dispatcher and receiver class using the value of the sum of rows (r) and the sum of column (c) as shown in Eq. 3.12 and Eq. 3.13. The critical factors acquiring the positive value of $(r-c)$ have higher influence and priority over other factors, therefore, referred to as dispatchers. However, negative values of $(r+c)$ have less influence on others, thus tend to get influenced from others, therefore, termed as a receiver. A higher value of $(r+c)$ shows the great degree of relationship among each factor.

$$S = (s_{i,j})_{m\times m}, \quad i, j \in \{1, 2, 3, ..., m-1, m\} \tag{3.12}$$

$$DP = \left(\sum_{j=1}^{n} s_{i,j}\right)_{n\times 1}, \quad RE = \left(\sum_{i=1}^{n} s_{i,j}\right)_{n\times 1} \tag{3.13}$$

5. Later, considering the factors that have a higher influence level than a certain threshold value in total relation matrix S are converted into a causality map referred to as impact digraphs.

$$\alpha = 1/m^2 \sum_{i=1}^{n}\sum_{j=1}^{n} s_{i,j} \tag{3.14}$$

The ranking of factors can be done on the basis of the absolute value of $r-c$ or the degree of relationship can be given through Eq. 3.15:

$$r(Bk) = \text{order } Bk(BP + C) \tag{3.15}$$

4. Experimental Result Analysis and Discussion. F-AHP method has been applied in various fields for solving complex decision problems. The structure of problems can be solved using numerous sub-factors using AHP, considering both qualitative and quantitative criteria in the same decision framework. The method is a multi-criteria technique based on selecting or ranking the best set of alternatives. Fuzzy AHP is a reliable tool for selecting significant factors, weight assignment to each identified factor and evaluation of alternative solutions in hierarchic matrices. To compute the relative weights, the fuzzy linguistic scale of importance has been considered as presented in Table 4.1. The opinions were collected by experts on the basis of their knowledge related to relevant factors and sub-factors to develop matrixes for pairwise comparison. The rating performed for each pair of factors important for embracing blockchain technology in human resource management practices is shown in Table 4.2. The AHP methodology consists of three principles: (i) Construction of hierarchical structure for the problem (ii) pair-wise comparative judgement on the basis of an expert's opinion that calculates the local weight (relative importance) on the identified criteria as shown in Table 4.3. (iii) synthesis of the local weight (relative importance) into global weights (Global importance) leading to the selection of the final decision as

Table 4.1: Fuzzy Linguistic Scale

Scale Importance	Fuzzy Scale	Fuzzy Scale Reciprocal
Equal	(1,1,1)	(1,1,1)
Moderate	(2,3,4)	(1/4,1/3,1/2)
Strong	(4,5,6)	(1/6,1/5,1/4)
Very Strong	(6,7,8)	(1/8,1/7,1/6)
Extremely Strong	(9,9,9)	(1/9,1/9,1/9)
Intermediate	(1,2,3) (3,4,5) (5,6,7) (7,8,9)	(1/3,1/2,1) (1/5,1/4,1/3) (1/7,1/6,2/5) (1/9,1/8,1/7)

Table 4.2: Factors value obtained from expert opinion

	Talent Acquisition	Payroll Processing	Compliance and Regulatory	Performance evaluation management
Talent Acquisition	1	9	4	4
Payroll Processing	0.111111	1	5	5
Compliance and Regulatory	0.25	0.2	1	7
Performance evaluation management	0.25	0.2	0.142857	1

Table 4.3: Factors value obtained from expert opinion

	Talent Acquisition	Payroll Processing	Compliance and Regulatory	Performance evaluation management
Talent Acquisition	1	9	4	4
Payroll Processing	0.111111	1	5	5
Compliance and Regulatory	0.25	0.2	1	7
Performance evaluation management	0.25	0.2	0.142857	1

Table 4.4: Fuzzy weights matrix for pair-wise comparison

Fuzzy AHP	Talent Acquisition	Payroll Processing	Compliance and Regulatory	Performance evaluation management		FUZZY Weights	
Talent Acquisition	1, 1, 1	9, 9, 9	3, 4, 5	3, 4, 5	0.45913	0.59572	0.76278
Payroll Processing	0.11,0.11,0.11	1, 1, 1	4, 5, 6	4, 5, 6	0.17671	0.22201	0.27853
Compliance and Regulatory	0.2, 0.25, 0.33	0.16,0.2,0.25	1,1,1	6,7,8	0.10233	0.13227	0.17796
Performance evaluation management	0.2, 0.25, 0.33	0.16, 0.2, 0.25	0.12, 0.14, 0.16	1, 1,1	0.03887	0.04999	0.06761

shown in Table 4.4. In spite of the distinct intellectual abilities of decision-makers, triangular fuzzy sets theory

Table 4.5: CTMC model parameters

Fuzzy AHP	Talent Acquisition	Payroll Processing	Compliance and Regulatory	Performance evaluation management	Normalized Values
Talent Acquisition	1, 1, 1	9, 9, 9	3, 4, 5	3, 4, 5	0.593232
Payroll Processing	0.11, 0.11, 0.11	1, 1, 1	4, 5, 6	4, 5, 6	0.221041
Compliance and Regulatory	0.2, 0.25, 0.33	0.16, 0.2, 0.25	1, 1, 1	6, 7, 8	0.134654
Performance evaluation management	0.2, 0.25, 0.33	0.16, 0.2, 0.25	0.12, 0.14, 0.16	1, 1, 1	0.051073

Table 4.6: Normalized fuzzy decision values

	B2	B3	B4	B5	B6	B6	B7	B8	B8	B9	B9
B2	0.020	0.137	0.176	0.098	0.098	0.098	0.059	0.078	0.078	0.078	0.078
B3	0.003	0.020	0.137	0.059	0.098	0.098	0.039	0.039	0.039	0.098	0.098
B4	0.002	0.002	0.020	0.098	0.137	0.137	0.059	0.039	0.039	0.039	0.039
B5	0.004	0.007	0.004	0.020	0.098	0.098	0.098	0.137	0.137	0.137	0.137
B6	0.004	0.004	0.003	0.004	0.020	0.020	0.020	0.039	0.039	0.039	0.039
B6	0.004	0.004	0.003	0.004	0.020	0.020	0.020	0.039	0.039	0.039	0.039
B7	0.007	0.010	0.007	0.004	0.020	0.020	0.020	0.059	0.059	0.078	0.078
B8	0.005	0.010	0.010	0.003	0.010	0.010	0.010	0.020	0.020	0.137	0.137
B8	0.005	0.010	0.010	0.003	0.010	0.010	0.007	0.020	0.020	0.137	0.137
B9	0.005	0.004	0.010	0.003	0.010	0.010	0.005	0.003	0.003	0.020	0.020
B9	0.005	0.004	0.010	0.003	0.010	0.010	0.005	0.003	0.003	0.020	0.020

Table 4.7: Normalized matrix of expert opinion

	B2	B3	B4	B5	B6	B6	B7	B8	B8	B9	B9	SUM
B2	0.027	0.152	0.215	0.137	0.174	0.174	0.105	0.140	0.140	0.191	0.191	1.645
B3	0.008	0.027	0.151	0.080	0.142	0.142	0.067	0.077	0.077	0.162	0.162	1.094
B4	0.006	0.008	0.028	0.107	0.165	0.165	0.082	0.077	0.077	0.101	0.101	0.916
B5	0.010	0.015	0.017	0.027	0.120	0.120	0.115	0.167	0.167	0.217	0.217	1.192
B6	0.006	0.007	0.007	0.007	0.026	0.026	0.024	0.047	0.047	0.061	0.061	0.317
B6	0.006	0.007	0.007	0.007	0.026	0.026	0.024	0.047	0.047	0.061	0.061	0.317
B7	0.009	0.014	0.014	0.009	0.030	0.030	0.026	0.069	0.069	0.110	0.110	0.490
B8	0.007	0.013	0.017	0.007	0.020	0.020	0.016	0.027	0.027	0.157	0.157	0.467
B8	0.007	0.013	0.017	0.007	0.019	0.019	0.012	0.027	0.027	0.157	0.157	0.462
B9	0.006	0.005	0.013	0.005	0.014	0.015	0.008	0.007	0.007	0.026	0.026	0.132
B9	0.006	0.005	0.013	0.005	0.014	0.015	0.008	0.007	0.007	0.026	0.026	0.132

is considered an appropriate technique to minimize uncertainty and biasness. The fuzzy AHP is preferential, due to its simplicity and higher consistency.

The proposed M-DEMATEL is a novel method that reflects the casual-effect relationship among the critical sub-factors by accommodating the perceptions of the respondents who are experts in the particular industry. The ranking obtained from fuzzy AHP shows that performance factors have been less significant as compared to other factors therefore, the normalized matrix considered for M-DEMATEL evaluation is shown in Table 4.5. Further, the normalized direct relationship matrix is computed as depicted in Table 4.6. The study presents the benefits of implementing blockchain in HRM activities calculated from the prominence score. The

Table 4.8: Normalized total direct relationship matrix

	B2	B3	B4	B5	B6	B6	B7	B8	B8	B9	B9
B2	0.016	0.092	0.131	0.083	0.106	0.106	0.064	0.085	0.085	0.116	0.116
B3	0.005	0.016	0.092	0.049	0.086	0.086	0.041	0.047	0.047	0.098	0.098
B4	0.004	0.005	0.017	0.065	0.100	0.100	0.050	0.047	0.047	0.061	0.061
B5	0.006	0.009	0.010	0.017	0.073	0.073	0.070	0.102	0.102	0.132	0.132
B6	0.003	0.004	0.004	0.004	0.016	0.016	0.015	0.028	0.028	0.037	0.037
B6	0.003	0.004	0.004	0.004	0.016	0.016	0.015	0.028	0.028	0.037	0.037
B7	0.005	0.009	0.009	0.005	0.018	0.018	0.016	0.042	0.042	0.067	0.067
B8	0.004	0.008	0.010	0.004	0.012	0.012	0.009	0.016	0.016	0.095	0.095
B8	0.004	0.008	0.010	0.004	0.012	0.012	0.007	0.016	0.016	0.095	0.095
B9	0.003	0.003	0.008	0.003	0.009	0.009	0.005	0.004	0.004	0.016	0.016
B9	0.003	0.003	0.008	0.003	0.009	0.009	0.005	0.004	0.004	0.016	0.016

Table 4.9: Representation of degree of influence

	B2	B3	B4	B5	B6	B6	B7	B8	B8	B9	B9	D	D+R	D-R
B2	0.016	0.092	0.131	0.083	0.106	0.106	0.064	0.085	0.085	0.116	0.116	1.000	1.058	0.942
B3	0.005	0.016	0.092	0.049	0.086	0.086	0.041	0.047	0.047	0.098	0.098	0.665	0.827	0.503
B4	0.004	0.005	0.017	0.065	0.100	0.100	0.050	0.047	0.047	0.061	0.061	0.556	0.860	0.253
B5	0.006	0.009	0.010	0.017	0.073	0.073	0.070	0.102	0.102	0.132	0.132	0.724	0.966	0.482
B6	0.003	0.004	0.004	0.004	0.016	0.016	0.015	0.028	0.028	0.037	0.037	0.193	0.648	-0.263
B6	0.003	0.004	0.004	0.004	0.016	0.016	0.015	0.028	0.028	0.037	0.037	0.193	0.648	-0.263
B7	0.005	0.009	0.009	0.005	0.018	0.018	0.016	0.042	0.042	0.067	0.067	0.298	0.594	0.002
B8	0.004	0.008	0.010	0.004	0.012	0.012	0.009	0.016	0.016	0.095	0.095	0.284	0.703	-0.136
B8	0.004	0.008	0.010	0.004	0.012	0.012	0.007	0.016	0.016	0.095	0.095	0.281	0.700	-0.139
B9	0.003	0.003	0.008	0.003	0.009	0.009	0.005	0.004	0.004	0.016	0.016	0.080	0.851	-0.690
B9	0.003	0.003	0.008	0.003	0.009	0.009	0.005	0.004	0.004	0.016	0.016	0.080	0.851	-0.690
R	0.058	0.162	0.303	0.242	0.456	0.456	0.296	0.419	0.419	0.770	0.770			

normalized total direct relationship matrix is evaluated and represented in Table 4.7. The normalized total relation matrix is computed by summing the direct effects and all the indirect effects. A better understanding of interrelationships can be obtained by thoroughly analyzing the total relation matrix as represented by S in equation 12.

D+R can be known as the degree of the central role that illustrates the strength that are given (Dispatched) and received (Receiver) of the factor. Similarly, the vertical axis vector (D-R) called "Relation" shows the net effect that the factor contributes to the system. If (D-R) is positive, then the factor can be categorized into cause groups; if (D+R) is negative, then the factor can be clustered into the effect group. Certainly, a Casual diagram can be generated by plotting the set of data (D+R, D-R), showing the reasonable observation for strategic decision making as shown in Table 4.8.

5. Managerial implication. The study finding has significant managerial implications and insights that will enable multinational companies to integrate blockchain into HRM functions. According to the literature review strategy, the applications of blockchain technology are eventually gaining alignment with the HRM sector. Therefore, this paper facilitates the organizational decision-makers, HR practitioners expert IT managers as

well as industry 4.0 developers to capture all appropriate information about how blockchain and HRM systems may work together to benefit both the company and the employees. The blockchain-based human resource recruitment system (BcHRS) will digitally verify official credentials and resumes to ease the HR professional's recruitment process in identifying the right candidate for an open position. The hiring process will undoubtedly transform with the added benefits of blockchain in terms of efficiency, quality, and low costs. Similarly, a global blockchain payroll solution can simply provide a speedier solution than current approaches. As a result of blockchain in payroll management, HR professionals or managers handling employee payroll, tax compliances, medical information, and vacation may find it easier to control cross-border expenses. Moreover, our findings include that blockchain technology offers a higher level of security than many existing technologies and also blockchain might also be utilized for employee learning records or other information stored in HRM systems. HR has the chance to improve workplace environments in accordance with corporate needs and increasing employee expectations. Blockchain is one of the technologies in which HRM should invest as part of its technological innovation in the HRM system.

6. Conclusion and Future Scope. Industry 4.0 is a paradigm that is transforming the way organizations operate by utilizing cutting-edge technology. The study was conducted to identify the critical success factors that encourage blockchain adoption in HRM and to establish the causal relationship among various sub-factors. The exploration to determine the significant factors as well subfactors of BCT adoption and its framework using a combined F-AHP and M-DEMATEL method is derived. The application of technology enabled solution in Industry 4.0 is estimated to rise and benefit a wide range of sectors. The result analysis indicates that the use of advance technology enhances and optimizes various HRM sectors to thrive in the digitization era for IOT smart cities.

In future, a blockchain-based intelligent HRM portal might give each employee as well as manager a single, targeted, and frequently customized access point. Employee self-service interaction system (ESS) and HR manager self-service system (MSS) are the two basic components and prerequisites for capturing those prospective applications. These two essential themes will be crucial in building blockchain-based smart HRM corporations. Further, analysis can be conducted for addressing the barriers and different dimensions of HRM to overcome the shortcomings while the adoption of intelligent solution.

REFERENCES

[1] L. Hughes, YK. Dwivedi, S. K. Misra, N. P. Rana, V. Raghavan, V. Akella,"Blockchain research, practice and policy: Applications, benefits, limitations, emerging research themes and research agenda," *International Journal of Information Management*, vol. 49, pp.114—119, 2019.

[2] Gartner, "Blockchain technology: what's ahead?" 2020.

[3] A. Hooper, D. Holtbr," Blockchain technology in international business: changing the agenda for global governance," *Review of International Business and Strategy*, vol. 30, pp. 183–200, 2020.

[4] S. Saxena,B. Bhushan, M. A. Ahad, "Blockchain based solutions to secure IoT: Background, integration trends and a way forward," *Journal of Network and Computer Applications*, vol. 181, pp. 103050, 2021.

[5] D.L. Stone, D. L . Deadrick, K. M. Lukaszewski, R. Johnson. "The influence of technology on the future of human resource management, " *Human resource management review* , vol. 25, pp. 216–231,no.2 2015.

[6] A. Pfeiffer, S. Bezzina, T. Wernbacher , V. Vella, A. Dingli, N. Denk, "The use of blockchain technologies to issue and verify micro-credentials for customised educational journeys: presentation of a demonstrator," *presentation of a demonstrator* , 2021.

[7] S. Secinaro, D. Calandra, P. Biancone, "Blockchain, trust, and trust accounting: can blockchain technology substitute trust created by intermediaries in trust accounting? A theoretical examination, " *International Journal of Management Practice* , vol. 14, no. 2, pp.129—145, 2021.

[8] L.P. Mahyuni, R. Adrian, G. S. Darma, N. N. K. Krisnawijaya, I. G. A. A. P. Dewi, G. P. L. Permana, "Mapping the potentials of blockchain in improving supply chain performance," *Cogent Business & Management*, vol. 7, no. 1, pp. 1788329, 2020.

[9] JP. Ashwini, HA. Sanjay, MC. Nayana, KA. Shastry KA, MM. MK, "Framework for Performance Enhancement of MPI based Application on Cloud," *Scalable Computing: Practice and Experience*, vol. 24, no. 1, pp. 55-67, 2023.

[10] D. Folkinshteyn, M. Lennon, "Braving Bitcoin: A technology acceptance model (TAM) analysis," *Journal of Information Technology Case and Application Research* , vol. 18, no. 4, pp. 220-249, 2016.

[11] M. D. Williams,N. P. Rana, Y. K. Dwivedi, "The unified theory of acceptance and use of technology (UTAUT): a literature review," *Journal of enterprise information management* , 2015

[12] C. S. Yu, "Factors affecting individuals to adopt mobile banking: Empirical evidence from the UTAUT model," *Journal of electronic commerce research* , vol. 13, no. 2, pp. 104, 2012.
[13] C. M. Chang, L. W. Liu, H. C. Huang, H. H. Hsieh, "Factors influencing online hotel booking: Extending UTAUT2 with age, gender, and experience as moderators," *Information* , vol. 10, no. 9, pp. 281, 2019.
[14] C. Morosan, A. DeFranco, "It's about time: Revisiting UTAUT2 to examine consumers' intentions to use NFC mobile payments in hotels," *International Journal of Hospitality Management*, vol. 53, pp. 17-29, 2016.
[15] V. Venkatesh, J. Y. Thong, X. Xu, "Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology," *MIS quarterly*, pp. 157-178, 2012.
[16] S. Olnes, J. Ubacht, M. Janssen, "Blockchain in government: Benefits and implications of distributed ledger technology for information sharing," *Government Information Quarterly*, vol. 34,, no. 3, pp.355-364, 2017.
[17] H. Albayati, S. K. Kim, J. J. Rho, "Accepting financial transactions using blockchain technology and cryptocurrency: A customer perspective approach," *Technology in Society* , vol. 62, pp. 10132, 2020.
[18] T. McGhin, K. K. R. Choo, C. Z. Liu, D. He, "Blockchain in healthcare applications: Research challenges and opportunities," *Journal of Network and Computer Applications* , vol. 135, pp. 62-75, 2019.
[19] S. Mansfield-Devine, "Beyond Bitcoin: using blockchain technology to provide assurance in the commercial world," *Computer Fraud & Security*, no. 5, pp. 14-18, 2018.
[20] Y. Chang, J. Xu, K.Z. Ghafoor, "An IOT and Blockchain Approach for the Smart Water Management System in Agriculture,"*Scalable Computing: Practice and Experience*, vol. 22, no. 2, pp. 105–116, 2021
[21] J. Xie, H. Tang, T. Huang, F. Yu, R. Xie, J. Liu, Y. Liu, A survey of blockchain technology applied to smart cities: Research issues and challenges," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 3, pp. 2794-2830, 2019.
[22] O. Ali, M. Ally, Y. Dwivedi, "The state of play of blockchain technology in the financial services sector: A systematic literature review," *International Journal of Information Management*, vol. 54, no. 9, pp. 10219, 2020.
[23] Tiago M. Fernandez-Carames, Fraga-Lamas. Paula, "A Review on the Use of Blockchain for the Internet of Things, " *IEEE Access*, no. 28426, pp. 32979-33001, 2018.
[24] S. Wan, M. Li, G. Liu, C. Wang, "Recent advances in consensus protocols for blockchain: a survey," *Wireless networks*, vol. 26, no. 8, pp. 5579-5593, 2020.
[25] Maciel M. Queiroz, Fosso Wamba Samuel, De Bourmont Marc, Telles Renato," Blockchain adoption in operations and supply chain management: empirical evidence from an emerging economy," *International Journal of Production Research*, vol. 59, no. 20, pp. 6087-6103, 2021.
[26] H. Treiblmaier, "The impact of the blockchain on the supply chain: a theory-based research framework and a call for action," *an international journal*, 2018.
[27] F. Casino, T.K. Dasaklis, C. Patsakis," A systematic literature review of blockchain-based applications: Current status, classification and open issues," *Telematics and informatics*, vol. 36, pp. 55-81, 2019.
[28] L. W. Wong, G. W. H. Tan, V. H. Lee, K. B. Ooi, A. Sohal," Unearthing the determinants of Blockchain adoption in supply chain management," *International Journal of Production Research*, vol. 58, no. 7, pp. 2100-2123, 2020.
[29] D. Tapscott, A. Tapscott, "The impact of the blockchain goes beyond financial services," *Harvard business review*, vol. 10, pp.7, 2016.
[30] R. Kumar, R. Tripathi, "Secure healthcare framework using blockchain and public key cryptography," *Blockchain Cybersecurity, Trust and Privacy ., Cham* , pp. 185-202, 2020.
[31] A. Dolgui, D. Ivanov, B. Sokolov, "Reconfigurable supply chain: The X-network," *International Journal of Production Research* , vol. 58, no . 13, pp. 4138-4163, 2020.
[32] I. Makhdoom, M. Abolhasan, H. Abbas, W. Ni, "Blockchain's adoption in IoT: The challenges, and a way forward," *Journal of Network and Computer Applications*, vol. 125, pp. 251-279, 2019.
[33] J. Mendling, I. Weber, W.V.D. Aalst, J.V. Brocke, C. Cabanillas, F. Daniel, "Blockchains for business process management-challenges and opportunities," . *ACM Transactions on Management Information Systems (TMIS)*, vol. 9,, no. 1, pp.1-16, 2018.
[34] P. Fairley, " The ridiculous amount of energy it takes to run Bitcoin," *IEEE Spectrum*, vol. 28, 2017.
[35] S. Saberi, M. Kouhizadeh, J. Sarkis, L. Shen," Blockchain technology and its relationships to sustainable supply chain management," *International Journal of Production Research* ,vol. 57, no. 7, pp. 2117-2135, 2019.
[36] A. Reyna, M. Cristian,C. Jaime, E. Soler, M. Diaz, "On blockchain and its integration with IoT. Challenges and opportunities," *Future generation computer systems*, vol. 88 pp.173-190, 2018.
[37] G. Dagher, J. Mohler, M. Milojkovic, P. Marella, "Ancile: Privacy-preserving framework for access control and interoperability of electronic health records using blockchain technology," *Sustainable cities and society* , vol. 39,pp.283-297, 2018.
[38] I. Britchenko, T. Cherniavska, B. Cherniavskyi, "Blockchain technology into the logistics supply," *Tarnobrzeg: Wydawnictwo Państwowej Wyższej Szkoły Zawodowej im. prof ...*, 2018.
[39] S. Brijesh, R. Pillai, "Technology and talent analytics for talent management–a game changer for organizational performance, " *International Journal of Organizational Analysis*, 2019.
[40] J. Frizzo-Barker, P. Chow-White, P. Adams, J. Mentanko, D. Ha, S. Green, "Blockchain as a disruptive technology for business: A systematic review," *International Journal of Information Management*, vol. 51,pp.102029, 2020.
[41] PV. Ingold, M. Langer, "Resume= Resume? The effects of blockchain, social media, and classical resumes on resume fraud and applicant reactions to resumes," *Computers in Human Behavior*, vol. 144, pp. 106573,2021.
[42] R. Arenas, P. Fernandez1717, "CredenceLedger: a permissioned blockchain for verifiable academic credentials," *2018 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC)*, pp. 1-6,2018.
[43] YC. Chen, H-J. Wu, C. Wang, C. Yeh, L. Lew, I-C. Tsai, "Applying blockchain technology to develop cross-domain digital talent," *2019 IEEE 11th International Conference on Engineering Education (ICEED)*, vol. 34,, no. 3, pp.113-117,

2019.
[44] H. Mishra, M. Venkatesan, "Blockchain in human resource management of organizations: an empirical assessment to gauge HR and non-HR perspective," *Journal of Organizational Change Management*,2021.
[45] E. Bertino, A. Bertino,Z. Sura "Data transparency with blockchain and AI ethics," *Journal of Data and Information Quality (JDIQ)*, vol. 11 pp.1-8, 2019.
[46] R. Kumar, R. Tripathi, "Large-scale data storage scheme in blockchain ledger using ipfs and nosql," *Large-Scale Data Streaming, Processing, and Blockchain Security*, pp. 91–116, 2021.
[47] Y. Nurhasanah, D. Prameswari, O. Fachrunnisa, "Blockchain-Based Solution for Effective Employee Management," *International conference on smart computing and cyber security: strategic foresight, security challenges and innovation* , pp. 147–154, 2020.
[48] EB. Sifah, H. Xia, C. N. A. Cobblah, Q. Xia, Gao, "BEMPAS: a decentralized employee performance assessment system based on blockchain for smart city governance," *IEEE Access*, vol. 8,, no. 1, pp.99528–99539, 2020.
[49] EB. Sifah, H. Xia, C. N. A. Cobblah, Q. Xia, Gao, "Human resource information management model based on blockchain technology," *2017 IEEE symposium on service-oriented system engineering (SOSE)*, pp. 168–173, 2017.
[50] R. Kumar, R. Tripathi, "Building an ipfs and blockchain-based decentralized storage model for medical imaging," *Advancements in Security and Privacy Initiatives for Multimedia Images*, pp.19–40, 2021.
[51] M. Velasquez, P.T. Hester, "An analysis of multi-criteria decision making methods," *International journal of operations research* , vol. 10,, no. 2, pp.56–66, 2013.
[52] T.P.N. Le, A. Genovese, L.S.C. Koh, others, "Using FAHP to determine the criteria for partner's selection within a green supply chain: The case of hand tool industry in Taiwan," *Journal of Manufacturing Technology Management*, 2012.
[53] C. Chauhan, A. Singh, S. Jharkharia, "An interpretive structural modeling (ISM) and decision-making trail and evaluation laboratory (DEMATEL) method approach for the analysis of barriers of waste recycling in India," *Journal of the Air & Waste Management Association*, vol. 68,, no. 2, pp.100-110, 2018.
[54] D.G. Bonett, A. T. Wright, "Cronbach's alpha reliability: Interval estimation, hypothesis testing, and sample size planning," *Journal of organizational behavior*, vol. 36,, no. 1, pp. 3–15, 2015.
[55] Y. Cheng, S. Zhou,, "Research of distribution center site selection based on fuzzy analytic hierarchy process," *Proceedings of the 22nd International Conference on Industrial Engineering and Engineering Management 2015*, pp.335-342, 2016.
[56] S. Lan, H. Zhang, R.Y. Zhon, GQ. Huang,, "A customer satisfaction evaluation model for logistics services using fuzzy analytic hierarchy process," *Industrial Management & Data Systems*, 2016.
[57] YC. Chen, H-J. Wu, C. Wang, C. Yeh, L. Lew, I-C. Tsai, "An integrated DEMATEL and Fuzzy ANP techniques for evaluation and selection of outsourcing provider for a telecommunication company," *Computers & Industrial Engineering*, vol. 86, pp. 137–146, 2015.

Edited by: Pradeep Kumar Singh
Received: Nov 30, 2022
Accepted: May 15, 2023

Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2023 SCPE. Volume 24, Issues 2, pp. 97–114, DOI 10.12694/scpe.v24i2.2117

AN EFFECTIVE DIABETIC RETINOPATHY DETECTION SYSTEM USING DEEP BELIEF NETS AND ADAPTIVE LEARNING IN CLOUD ENVIRONMENT

PRAVEEN MODI* AND YUGAL KUMAR†

Abstract. The major reason behind the blindness of the diabetes patients is diabetic retinopathy. It can be characterized as an eye disease that affects the retina of eye due to diabetes mellitus. The detection of diabetic retinopathy in early stage is a challenging task to ophthalmologists. This paper presents a diabetic retinopathy detection system for accurate detection of DR in the patients. The proposed diabetic retinopathy detection system is the combination of several preprocessing technique and deep belief nets. The aim of preprocessing technique is to enhance the images, edge detection, and segmentation. Further, the deep belief nets are adopted for the accurate detection of DR. But, the parameter tuning of weight, bias and learning rate have significant impact on the performance of deep belief nets. This work also addresses these issues of deep belief nets though an adaptive learning strategy for learning rate and updated mechanism for weight and bias issues. The proposed system is implemented in cloud environment. It is utilized to store the information regarding DR and communication between doctors and patients. Further, the efficacy of the proposed diabetic retinopathy detection system is tested over an image dataset and it comprises of three thousand two hundred eye images include with diabetes retinopathy and no diabetes retinopathy. The results are evaluated using accuracy, sensitivity, specificity, F1-Score and AUC parameters. The results of proposed system are compared with KNN, SVM, ANN, InceptionV3, VGG16 and VGG19 techniques. The results showed that proposed diabetic retinopathy detection system obtains 91.28% of accuracy, 93.46% of sensitivity, 94.84 of specificity and 94.14 of F1-Score rates than other techniques using 10-cross fold validation method. Hence, it is stated that proposed system detects diabetes retinopathy more accurate than other techniques.

Key words: Adaptive learning, Deep Belief Nets, Image, Diagnosis, Diabetes, Diabetic Retinopathy

1. Introduction. Diabetes mellitus (DM) can be characterized as chronic disease throughout worldwide and fourth foremost reason of people death [28]. A recent study showed that presently in the world, 336 million people are affected form diabetes mellitus and 7.7% more people will be affected with diabetes mellitus up to 2030 [33]. Further, the diabetic retinopathy (DR) can be defined as condition of DM that is responsible for blindness in diabetes patients [36]. It is related with type-1 and type-2 diabetes. The symptoms of DR can be visible for type-1 diabetes patients after fifteen years of diabetes, and 75%-90% of patients are suffering from DR symptoms. While, 60% of type-2 diabetes patients having symptoms of DR that are affected with diabetes more than 16 years. The contribution of DR into eye related disease up to 80%, especially patients that are diagnosed eye related disease more than ten years [1, 25]. The main points regarding the incremental growth of DR are lack of prior symptoms of DR, severely vision loss and untimely diagnosis. However, the existence of DR can be reduced with timely diagnosis and screening, proper medical treatment, and medication. The initial treatment for DR can be described in terms of fundus photography through ophthalmoscopy and grading the fundus images. Further, these retinal fundus images are examined by ophthalmologists for detecting the DR in manual order. The ophthalmologists examine the presence of cotton wool spots, retinal swellings, and hemorrhages [9]. Sometimes, the process of capturing the retinal fundus images can reveal irrelevant illuminations, blurred and darkened candidate regions, and non-uniformity of light distribution [38]. But, the high precision is required for detecting the DR, otherwise it will lead to the wrong decision and responsible for serious problems and sometimes permanent blindness in patients. If the image obtained through the fundoscopic test is highly saturated, then it is quite a tough task for ophthalmologists to proper visual investigation of diabetic retinopathy. It is also found that the occurrence of non-uniform illuminations can lead to bias prediction of DR [30]. Hence in pre-processing task, luminosity normalization is one of important aspect for generating the diverse set of retinal images [7].

*Department of Computer Science and Engineering and Information Technology, Jaypee University of Information Technology, Waknaghat, Solan, Himachal Pradesh, India (modi.240289@gmail.com)
†Department of Computer Science and Engineering and Information Technology, Jaypee University of Information Technology, Waknaghat, Solan, Himachal Pradesh, India (yugalkumar.14@gmail.com)

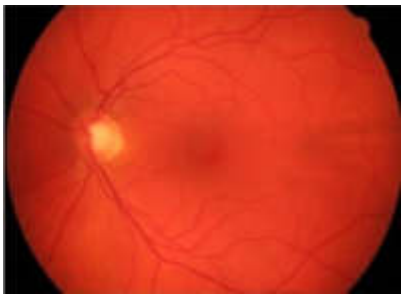


Fig. 1.1: Depicts the healthy retinal fundus images

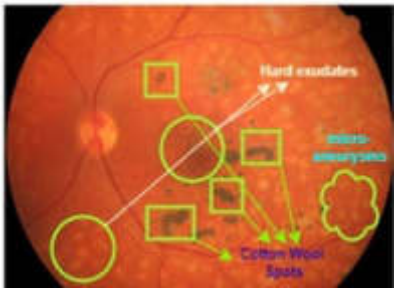


Fig. 1.2: Depicts diabetic retinal fundus images with retina complications

Fig. 1.1 shows the healthy retinal fundus image and Fig. 1.2 shows the diabetic retinal fundus images with retinal complications. Through the literature on diabetic retinopathy, it is found that diabetic retinopathy is divided into two categories- (i) binary classification of DR disease, and (ii) multiclass classification of DR disease. In literature, several techniques have been presented for accurate detection of DR disease based on eye fundus images. These techniques describe the candidate regions in terms of exudates, cotton wool spots, blood vessels, micro aneurysms and hemorrhages etc. Further, these techniques also consider the uniform features and extract the features manually [31, 11]. In turn, the detection rate of DR is not up to significant level.

1.1. Motivation and Contribution of the Work. This subsection presents the motivation and contribution of the work. Through literature, it is identified that CNN, MLP and several popular ML algorithms have been explored to find the more accurateand earlier detection of DR disease [35]-[37].However, the problems of non-uniform reflectance, over fitting, over saturation and biased solution are still remaining.Hence, the main focus of this work is to handle aforementioned issues related with accurate detection of DR disease based on eye fundus images.This work presents an automatic diabetic retinopathy detection system for accurate prediction of DR. The proposed system comprises of various pre-processing technique for handling the non-uniform reflectance and over saturation issues of fundus images.Further, the deep belief nets are integrated into proposed detection model as predictive classifier. It is also noticed that the performance of the deep belief nets is dependent on learning rate, weight and bias parameters. If, learning rate is not optimal, then it can not achieve contrastive divergence as well as optimal training of the model. Hence, to alleviate this issue, an adaptive learning strategy is incorporated into beep belief nets. The main contributions of the work are highlighted as below.

- To developed an automatic diabetic retinopathy detection system for accurate detection of DR disease through eye retinal fundus images.
- To handle the issue of non-uniform illumination and color of fundus images through threshold and luminosity function through processing of image step in the proposed diabetic retinopathy detection system.The color component(R,G,B) are extracted from the images and threshold function is designed to improve the RGB. The distribution of light is also adjusted through luminosity function.

- The statistical and spatial features are computed using the histogram and edge detection techniques. Further, k-means based segmentation technique is utilized for obtaining the appropriate clusters.
- A deep belief nets with adaptive learning strategy is integrated into proposed diabetic retinopathy detection system for accurate detection of DR disease.
- The performance of the proposed detection model is tested over fundus image dataset comprising of 3200 images and results are evaluated using accuracy, AUC, sensitivity, specificity and F1-Score parameters.
- The results of proposed model are compared with Incpetion V3, VGG16, VGG19, ANN, SVM,and KNN techniques. The efficacy of simulation results are also assessed though training/testing (70%-30%), training/testing (80%-20%), 5-cross fold validation and 10-cross fold validation methods.The results revealed that proposed diabetic retinopathy detection system is efficient to recognize the DR disease.

2. Related Works. This section summarizes the recent work reported on diabetic retinopathy diagnosis. Gao et al.[8] considered the diabetic retinopathy as one the major concern among the diabetes affected patients and it is also responsible for the blindness in diabetes. In this work, an automated diagnosis model is presented for prediction of diabetic retinopathy and this system is also capable for providing the suggestions to diabetic retinopathy patients. In the proposed system, deep convolutional neural network is utilized for prediction of DR. Furthermore, a diabetic retinopathy dataset is constructed using the retinal fundus images and severities of DR is divided into four classes. The performance of the proposed model is evaluated using accuracy parameter and results showed that proposed model achieves 88.72% of accuracy rate. The proposed model is also deployed into several hospitals and authors claimed that proposed model obtains 91.8% of accuracy with ophthalmologists.

Kandhasamy et al. [17] developed a diagnosis system for finding the severity of DR. The proposed diagnosis system is the combination of multi-level segmentation and SVM with GA. Some morphological operations are also applied to determine the clusters in the retinal images. The multi-level segmentation is employed on these clusters for extracting features. Further, a local binary pattern is also adopted for extracting the texture features of retina images. The effectiveness of the diagnosis system is evaluated using accuracy, sensitivity and specificity parameters. It is found that proposed diagnosis system achieves 99.3% accuracy rate, 97.14% of sensitivity rate and 100% specificity rate.

Qureshi et al. [22] presented an automatic recognition model of DR severity. The proposed recognition model is based on multilayer architecture of active learning, called ADL. In ADL, CNN is adopted for computing the features in automatic manner. Further, an excepted gradient length method is applied as active learning method for multilayer architecture, called ADL-CNN. The proposed ADL-CNN is in two folds– (i) consider the informative patches and ground truth label of images to train the model, and (ii) prognostication the image into five severity-levels of diabetic retinopathy. The results of the ADL-CNN model are evaluated using accuracy, sensitivity, specificity and F1-Score. It is noticed that proposed ADL-CNN achieves higher results than other methods in terms of accuracy, sensitivity, specificity and F1-Score.

Qiao et al. [21] considered the deep convolutional neural network to determine the presence of micro aneurysm in fundus. Further, GPUs are adopted for acceleration of CNN. In this study, fundus images are segmented using semantic segmentation algorithm and this algorithm divides the images in binary class- (i) normal, and (ii) infected. It is seen that CNN predicts the diabetic retinopathy as early NPDR, moderate NPDR, and severe NPDR. The performance of the CNN model is evaluated through sensitivity, specificity and average accuracy. It is seen that proposed CNN based model achieves higher accuracy rate.

Jebaseeli et al. [16] designed an IOT based Sustainable Diabetic Retinopathy Diagnosis System for effective treatment of DR. In the proposed system, the glucose level of diabetes patients is collected through Dexcom G4 Platinum sensors. After classification of diabetes, the retinal fundus images of patients are captured through smart camera. Further, the segmentation is performed on the captured images and a modified fuzzy c-means algorithm is employed for predicting the diabetic retinopathy. The results showed that IOT based diabetic retinopathy diagnosis system obtain higher accuracy, sensitivity and specificity rates than other existing models.

Mansour [19] developed CAD based diabetic retinopathy model for early detection and diagnosis of DR disease. The proposed retinopathy model contains an Alexnet DNN technique which is a variant of CNN to find the optimal solution for DR. The Gaussian mixture model is adopted to determine the region segmentation.

LDA based feature selection technique is employed for computing relevant features. It is observed that proposed retinopathy model obtains more than 0.7% classification accuracy rate. With spatial features, the proposed model having 94.4% accuracy rate.

Kaushik et al. [18] considered the reflectance properties issue of the fundus images as one of potential issue for accurate detection of diabetic retinopathy. In this work, gray world color constancy algorithm is utilized for the luminosity normalization of the images. The fundus images are diagnosed using stacked deep learning technique. The performance of the Stacked DNN is assessed over peak signal to noise ratio (PSNR), mean squared error (MSE), accuracy, F-measure, sensitivity, specificity, recall and precision. The simulation results showed that Stacked DNN obtains 87.45% accuracy rate for multiclass classification. To determine the abnormality in retinal images,

Hemanth et al. [13] presented a modified Hopfield neural network (MHNN) for diagnosis of diabetic retinopathy. In modified Hopfield neural network, weights are optimized using an updated mechanism and changed in each iteration instead of fixed weights. The novelty of the proposed MHNN is tested over five hundred forty retinal images and results are evaluated using sensitivity, specificity and accuracy parameters. The MHNN obtains 99.25% of accuracy rate than HNN. In continuation of their work,

Hemanth et al. [14] presented a hybrid method for improving the diagnosis rate of diabetic retinopathy. The hybrid method is the combination of the image processing and deep learning method. The histogram equalization method is utilized for image processing. For the prediction task, convolutional neural network technique is adopted. The efficacy of the hybrid method is investigated over four hundred retinal images taken from MESSIDOR database. The results are evaluated using accuracy, sensitivity, specificity, F-Score and G-mean. It is noticed that hybrid model provides 97% of accuracy rate than existing methods. To alleviate the problems of misdiagnosis, reducing time, cost and effort,

Alyoubi et al. [2] developed two deep learning based model for effective diagnosis of diabetic retinopathy. These models are CNN512 and YOLOv3. The proposed model divides the retinal fundus image dataset into five severity class. Further, CNN512 considers the entire image as an input and predicts into five stages such as no-DR, mild, moderate, severe and proliferative DR. CNN512 model obtains 84.1% of accuracy rate. YOLOv3 model is adopted for detecting and localizing the DR lesions. Finally, both models are integrated into a single model and it is seen that combination of both models achieves higher accuracy rate of 89%.

Sharafeldeen et al. [24] designed a computer assisted diagnostic model for early detection of diabetic retinopathy. This work considers the optical coherence tomography scan for detecting DR. In the proposed diagnostic model, the segmentation approach is adopted for separating the retinal layers. Further, morphological and reflective markers are extracted from each layer and cumulative distribution function is applied for extracting the image driven markers. The SVM with linear kernel is utilized for diagnosis of diabetic retinopathy at each layer. The novelty of diagnostic model is examined over two hundred sixty OCT images and experimental results are evaluated using sensitivity, specificity, F1-score, and accuracy parameters. The experimental results showed that 96.15% of sensitivity, 99.23% of specificity, 97.66% of accuracy, and 97.69% of F1-score rates. Most of the works reported in literature on diabetic retinopathy are considered the high resolution images.

Wang et al. [32] considers the low resolution retinal fundus images for detecting the diabetic retinopathy. This work adopts the CNN technique to joint learning of multi-level tasks for grading of the DR, named it DeepMT-DR. The aim of DeepMT-DR is to handle low-level task of ISR, mid-level task of lesion segmentation and high-level task of disease severity classification. The efficiency of DeepMT-DR are tested over three image datasets and experimental results are evaluated using accuracy parameter. It is found that DeepMT-DR obtains 83.6% of accuracy rate.

Skouta et al. [26] designed an automated method for screening of diabetic retinopathy. The proposed automated method consists of modified CNN UNet architecture to determine retinal hemorrhages in fundus images. In this work, IDRiD dataset is adopted for evaluating the efficacy of automated model. In automated model, UNet is applied for training task and detecting the possible symptoms of DR. The experimental results are assessed through sensitivity, specificity and accuracy parameters. It is found that the proposed automated model achieves 80.49% of sensitivity, 99.68% of specificity, and 98.68% of accuracy rates. The selections of relevant features have significant impact on the performance of the classifiers especially with image dataset.

Vijayan et al. [29] considered the feature selection issue of image dataset and presented the simple color

histogram filter as feature selection method for retinal fundus images. KNN and J48 techniques are adopted for diagnosis of the diabetic retinopathy in fundus images. The performance of the proposed feature selection technique with J48 and KNN is evaluated using accuracy and ROC parameters. The results showed that KNN with feature selection method obtains 81.99% accuracy rate.

Gurcan et al.[12] presented an automated system for identification of diabetic retinopathy. The proposed system considers the retinal images for detection of diabetic retinopathy. The features are extracted from retinal images based on InceptionV3 including the transfer learning. The simulated annealing is utilized for selecting the relevant features. Finally, the diabetic retinopathy is detected using the XGBoost technique. The performance of the proposed automated model is evaluated using Messidor-II dataset. The results showed that proposed automated model obtains more than 92% of accuracy rate.

Gundluru et al.[10] considered the several issues like feature selection, optimization etc., of the existing diabetic retinopathy systems and presented a deep learning model for accurate detection of diabetic retinopathy. In the proposed model, PCA method is utilized for dimensionality reduction of the given dataset. While, harris hawks optimization algorithm is adopted for optimizing the feature extraction and classification processes. The proposed model is implemented on Diabetic Retinopathy Debrecen dataset and simulation results are evaluated using specificity, precision, accuracy, and recall parameters. The findings stated that proposed model obtains satisfactory results compared to existing systems.

Yue et al. [34]considered the limited capability to extract lesion-aware information and manual lesion annotations issues of diabetic retinopathy process. To handle aforementioned issues of DR, an end-to-end Attention-Driven Cascaded Network (ADCNet) is presented for grading of the diabetic retinopathy. In the proposed cascade network, the lesion-aware information is extracted using the hybrid attention module at shallow layer. The hybrid attention module is the combination of the multi-branch spatial attention and a loss-based attention. An aggregation strategy based on attention driven is also designed to get relevant features for DR. The APTOS and EyePACS datasets are chosen for examining the performance of the proposed cascade network. The results showed that proposed cascade network provides superior results than existing state of art methods.

To reduce the error rate and computational time, Chandran et al. [6] designed an auto-metric graph neural network (AGNN) for grading the diabetic retinopathy. Further, the noise in the images are eliminated throughAPPDRC filtering method. The features are extracted from retinal images based on GLCM based method. The weight of the AGNN is optimized using Capuchin search optimization algorithm. The performance of the AGNN model is evaluated using two popular eye retinal datasets i.e. ISBI 2018 and Messidor based on f-measure, execution time and accuracy metrics. The results showed that proposed model obtains more accurate results with both of datasets.

Canayaz [4]applied the wrapper method for selecting the more relevant features to detect the diabetic retinopathy. In their work, author explore the five hundred twelve features by using the EfficientNet and DenseNet models. Finally, two hundred fifty features are extracted based on the wrapper method. These method comprises of several popular meta-heuristic algorithms like binary bat algorithm (BBA), equilibrium optimizer (EO), gravity search algorithm (GSA), and gray wolf optimizer (GWO). The efficacy of the model is evaluated using the APTOS dataset based on the accuracy and kappa parameters. It is observed that proposed model obtains higher accuracy (96.32%) and kappa rate (0.98) compared to other algorithms.

To improve the detection accuracy of diabetic retinopathy Ragab et al. [23] developed a deep learning enable computer diagnosis model, called MDL-CADDR. In the proposed MDL-CADDR model, the image quality is enhanced in the pre-processing phase by applying filter and image contrast. Further, the region of interest is determined using Archimedes optimization algorithm (AAO) and relevant features are chosen based on Chimp Optimization Algorithm with DenseNet. Finally, detection of diabetes retinopathy is accomplished through skipping neural network. The well-known MESSIDOR dataset is utilized for evaluating the efficiency of MDL-CADDR model. The findings stated that MDL-CADDR model obtains the higher accuracy rate compared to other existing diabetic retinopathy model.

To early detection and diagnosis of diabetic retinopathy. Modi and Kumar[20] presented and efficient model for accurate detection of DR. The proposed model comprises of reprocessing, feature extraction, feature selection and classification. The noise are eliminated using the median filter and images are enhanced using

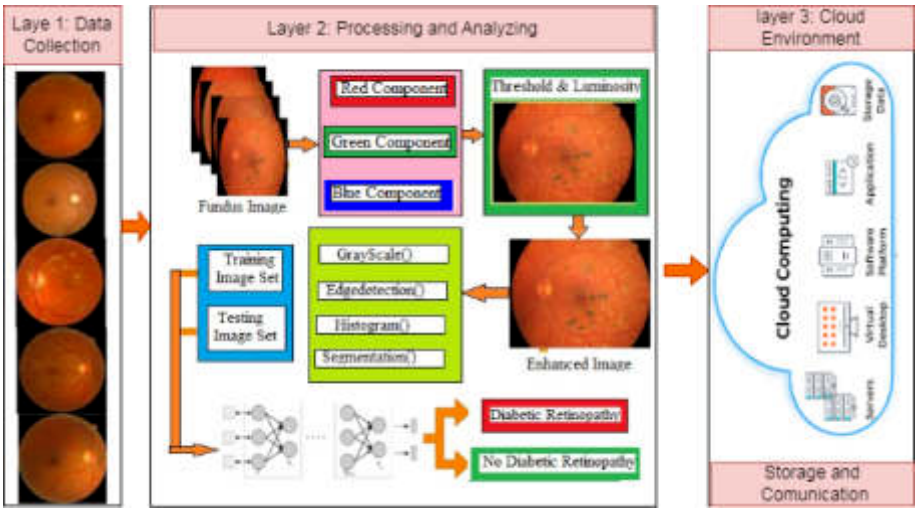


Fig. 3.1: Schematic Diagram of proposed diabetic retinopathy detection system

CLAHE algorithm. The k-Mean based segmentation technique is utilized for extracting the lesion region. The features are extracted using multi grained scanning technique, while relevant features are identified using bat based optimization algorithm. The classification task is performed using the DeepForest cascade technique. The efficacy of the proposed model is evaluated using two well-known diabetic retinopathy datasets based on accuracy, F-score, sensitivity and specificity. The results confirmed that proposed model obtains superior results for diabetic retinopathy compared to existing models. Table 2.1 summarizes the existing work of diabetic retinopathy detection in terms of segmentation method, feature selection method, classification model and potential parameters.

3. Proposed Diabetic Retinopathy Detection System. This section presents the proposed diabetic retinopathy detection system. The proposed diabetic retinopathy detection model is illustrated into Fig 3.1. The proposed detection system comprises of three layer-(i) data collection layer, (ii) processing and analyzing layer, (iii) cloud environment. The data collection layer is responsible for collection of data and working of this layer is explained in subsection 3.1. The second layer is the processing and analyzing Layer. The main activities of this layer is to perform the pre-processing activities, feature creation, and detection of diabetic retinopathy. The working of this layer is discussed in subsection 3.2. The third layer of the proposed model corresponds to cloud environment and working of this layer is discussed in subsection 3.3.

3.1. Data Collection Layer. This layer is responsible to collect the data regarding the diabetic retinopathy. In this work, a total a total of three thousand two hundred retinal images are considered for evaluating the performance of the proposed detection system. These images are downloaded from IEEE data portal (https://ieee-dataport.org). Further, two hundred twenty-four images are related to diabetic retinopathy symptoms and rest of the images have not any DR symptoms.

3.2. Processing and Analyzing Layer. The task of this layer is to process the images and analyze the performance of the proposed DR detection system. The image quality is enhanced based on resizing, extracts the feature and perform segmentation. The occurrence of DR is detected using adaptive deep belief network and performance of model is analyzes using well defined performance measures.

3.2.1. Processing of Image. Initially, the colored images are downloaded, stored in the folder on local system. The first step corresponds to enhance the retinal fundus images as processing of the images. In this step, the R, G, and B component of the images are extracted and further, threshold and luminosity mechanism are adopted for enhancing the quality of images.In turn, an enhanced quality images are obtained that contains high luminance and information instead of weak luminance and little information. In the next step, imresize()

Table 2.1: Summary of recent survey papers on diabetic retinopathy

Author	Segmentation Method	Feature Selection Method	Classification Method	Potential Parameter
Gao et al.[8]	*	*	Deep CNN	Detection Accuracy
Kandhasamy et al.[17]	*	Multi-level Segmentation and LBP	SVM Classifier	Sensitivity, Specificity and Accuracy
Qureshi et al.[22]	*	*	ADL	Sensitivity, Specificity, F1-Score and Accuracy
Qiao et al.[21]	Semantic Segmentation Method	MSSIM Maximization	CNN	Sensitivity, Specificity and Average Accuracy
Jebaseeli et al.[16]	*	Dexcom G4 Platinum sensors	Modified Fuzzy C-Means Algorithm	Sensitivity, Specificity and Accuracy
Mansour[19]	E-GMM	PCA & LDA	Alexnet DNN technique	Accuracy, Specificity, and Sensitivity
Kaushik et al.[18]	Gray World Color Constancy Algorithm	Gaussian Convolutional Deep Belief Network with Dwarf Mongoose Optimization Algorithm	Stacked DNN	Accuracy, Specificity, and Sensitivity
Hemanth et al.[13]	*	CNN	Modified Hopfield Neural Network	Sensitivity, Specificity and Accuracy
Hemanth et al.[14]	*	*	CNN	Specificity, Precision, Accuracy, and Recall
Alyoubi et al.[2]	YOLO V3	PCA	CNN512 and YOLO V3	Accuracy, Specificity, and Sensitivity
Sharafeldeen et al.[24]	Fuzzy C-means Clustering	*	SVM With Linear Kernel	Sensitivity, Specificity, F1-Score and Accuracy
Wang et al.[32]	*	*	CNN	Specificity, Precision, Accuracy, and Recall
Skouta et al.[26]	Semantic Segmentation Method	*	CNN UNet	Sensitivity, Specificity, F1-Score and Accuracy
Vijayan et al.[29]	*	*	KNN and J48	Accuracy and ROC
Gurcan et al.[12]	*	Simulated Annealing	InceptionV3, XGBoost	Accuracy
Gundluru et al.[10]	*	*	DNN-PCA-HHO	Sensitivity, Specificity, Recall, Precision
Yue et al.[34]	*	HAM	ADCNet	Specificity, Precision, Accuracy, and Recall
Chandran et al.[6]	*	*	AGNN Optimized with Capuchin Search Optimization	Accuracy, ROC
Canayaz[4]	*	GSA & GWO	SVM	Accuracy and Kappa-Score
Ragab et al.[23]	Archimedes Optimization	Chimp Optimization Algorithm	Spiking Neural Network	Accuracy
Modi and Kumar[20]	K-Mean	BAT Optimization Algorithm	Deep- Forest Cascading Technique	Accuracy, Precision, Recall, and Sensitivity

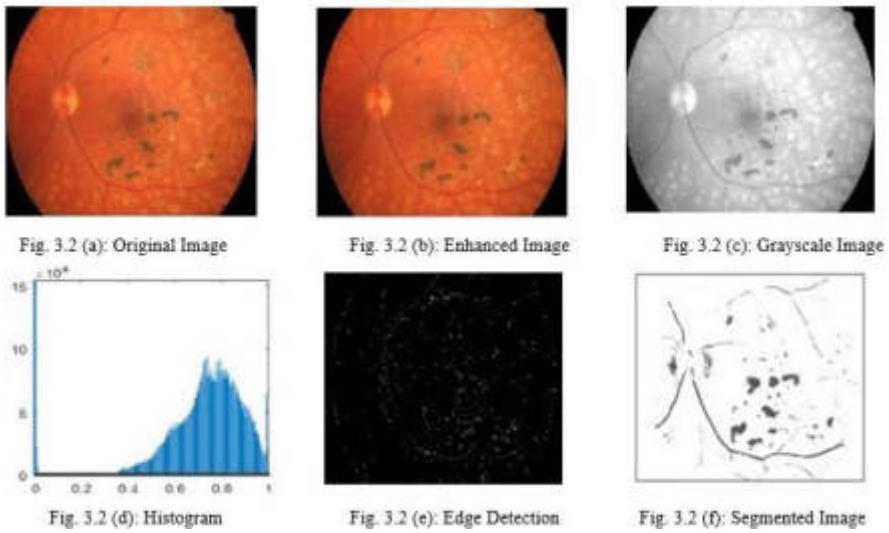


Fig. 3.2: (a-f) Depicts the processing of the image, Fig.3.1(a) shows the original retinal fundus image, Fig.3.1(b) illustrates the enhanced image in terms of RGB component of original image, Fig.3.1(c) shows the grayscale image of enhanced image, Fig.3.1(d) shows the histogram of corresponding grayscale image, Fig.3.1(e) shows the edge detection of grayscale image and Fig.3.1(f) shows the segments/clusters using CLAHE.

function is applied for resizing the images into 64×64 size and further, the RGB images are converted into grayscale as these images are represented through 8 bits and having pixel information in the range of 0 to 255. Further, it is observed that retinal layer is not of gray color and conversion into grayscale can be used to examine the image quality in normalization phase. This work considers the statistical and spatial features of detection of diabetic retinopathy. So, a hist() function is applied on the image for computing the statistical feature. Further, the edge of the grayscale images is detected through canny() function. The aim of edge detection is to determine the feature of images and check whether there are significant changes in the feature during the grayscale and also computing spatial features. Finally, the image segmentation is performed on grayscale images through imsegkmeans() function and number of K is set to 2. The aforementioned process is illustrated in the Fig 3.2 (a-f).

3.2.2. Deep Belief Nets With Adaptative Learning Strategy. Deep Belief Nets can be described as an improved variant of deep neural network [15] [3]. It contains multiple restricted Boltzmann machines (RBMs) which are stacked to each other. Further, the greedy approach is utilized for training of the stacked RBMs, and the stacked RBMs network is connected to a deep layer for final outcome, called deep belief nets. In DBN, an RMB is designed on the basis of two sequential hidden layer and input layer is characterized as output layer of the previous RBM, called visible layer. The distribution probability of visible input layer (VI) and hidden layers (HI) is described into equation (3.1a and 3.1b) and the hidden layers (HI) can be given as $HI_a = \in (a = 1, 2, 3....n)$ $HI_0 =$ VI.

$$DistProb(HI_1, HI_2, ..., HI_n/VI) = Prob\left(\frac{HI_n}{HI_{n-1}}\right) Prob\left(\frac{HI_{n-1}}{HI_{n-2}}\right) ...Prob\left(\frac{HI_1}{VI}\right) \tag{3.1a}$$

$$= \Pi_{a=1}^{n} Prob\left(\frac{HI_a}{HI_{a-1}}\right) \tag{3.1b}$$

The probability of bottom Prob(HI_1/ VI)with respect to visible layer (VI) and hidden layer (HI_a) is given using equation (3.2).

$$Prob(HI_a/HI_{a-1}) = \sigma\left(bi_j^a + \sum_{j=1}^{m} W_{i,j}^a \times Prob_j^{a-1}\right) \tag{3.2}$$

The probability of top inference (Prob($HI_a/HI_(a-1)$)) with respect to visible layer (VI) and hidden layer (HI_a) is given using equation (3.3).

$$Prob(HI_a/HI_{a-1}) = \sigma\left(bs_j^{a-1} + \sum_{j=1}^{m} W_{i,j}^{a-1} \times Prob_j^{a}\right) \tag{3.3}$$

Prior to precede the training of the deep belief nets, firstly, the concept of RBM is discussed as it can be worked like baseline method for deep belief nets. RBM can be defined as restricted Boltzmann machine (RBM) an extension of generative neural network [5]. Furthermore, it trains the network through the probability distribution function on given input set. RBM is more precisely described in terms of input layer and hidden layers. The hidden layers are utilized for learning the input data. An RBM consists of visible layer (VI) and hidden layer (HI) with values belong to 0, 1. A weight matrix ($W_{p\times q}$) can be defined to store the weight of connection among visible layer (VI) and hidden layer (HI) and the weight of j^{th} hidden layer (HI) and I^{th} visible layer (VI) is denoted as $((w)_{(i,j)})$. Further, the bias of the j^{th} hidden and the i^{th} visible layers are also computed to minimize the error and can be represents as bi_j and bs_i respectively. An energy function among the j^{th} hidden and the i^{th} visible layers is computed using equation (3.4).

$$ef(VI, HI) = \sum_{i=1}^{k} bs_i VI_i - \sum_{j=1}^{m} bi_j HI_j - \sum_{i=1}^{k}\sum_{j=1}^{m} W_{i,j} \times VI_i \times HI_j \tag{3.4}$$

In equation (3.4), $bs_i \in (i = 1, 2, \ldots, k)$ denotes the bias related visible layer (VI), $bi_j \in (j = 1, 2, \ldots, m)$ denotes the bias related to hidden layer (HI) and $w_{(i,j)} \in (i = 1, 2, \ldots, k$ and $j = 1, 2, \ldots, m)$. These are the learning parameters of RBM, while k and m denotes the number of features in visible and hidden layers. Hence, the configuration probability of visible and hidden layers is described using equation (3.5).

$$Prob(VI, HI) = \frac{1}{M} e^{-ef(VI,HI)} \tag{3.5}$$

In equation (3.5), M can be described as normalization term and it can be computed using equation (3.6).

$$M = \sum_{HI,VI} e^{-ef(VI,HI)} \tag{3.6}$$

Finally, the visible layer probability can be described in terms of hidden layer and it is computed using equation (3.7)

$$Prob(VI) = \frac{1}{M} \sum_{HI} e^{-ef(VI,HI)} \tag{3.7}$$

The independence probability of data distribution is easier to compute as neurons on same layer are not connected to each other. Hence, for randomly chosen neuron on visible layer (VI), the probability of j^{th} hidden layer (HI_j) is computed using equation (3.8).

$$Prob\left(HI_j = \frac{1}{VI}\right) = \sigma\left(bi_j + \sum_{i=1}^{k} W_{i,j} \times VI_j\right) \tag{3.8}$$

In equation (3.8), σ can be defined as linear sigmoid function. In similar manner, the probability of i^{th} visible layer (VI_i) with respect to hidden layer (HI) can be computed using equation (3.9).

$$Prob\left(VI_i = \frac{1}{HI}\right) = \sigma\left(bs_i + \sum_{j=1}^{m} W_{i,j} \times HI_j\right) \tag{3.9}$$

Further, a threshold function (ω) is employed for assessing the probability of hidden and visible layers. The computed probability is compared with threshold function(ω) as summarized into equations (3.10-3.11).

$$HI_j = \begin{cases} 1, \text{if } Prob(HI_j = 1/VI) > \omega \\ 0, \text{otherwise} \end{cases} \tag{3.10}$$

$$VI_i = \begin{cases} 1, \text{if } Prob(VI_i = 1/HI) < \omega \\ 0, \text{otherwise} \end{cases} \tag{3.11}$$

The probability of training can be enhanced through optimized values of weight and bias parameters. The weight and bias are updated using equations (3.12-3.14).

$$\triangle W_{k,l} = \varphi(\langle VI_{j,k}, HI_{j,m}\rangle_{data} - \langle VI_{j,k}, HI_{j,m}\rangle_{model}) \tag{3.12}$$

$$\triangle a_k = \varphi(\langle VI_{j,k}\rangle_{data} - \langle VI_{j,k}\rangle_{model}) \tag{3.13}$$

$$\triangle b_1 = \varphi(\langle HI_{j,m}\rangle_{data} - \langle HI_{j,m}\rangle_{model}) \tag{3.14}$$

3.2.3. Adaptive Learning Strategy. It is found that presence of multiple RBMs in deep belief nets architecture make the training process tedious and complex. In turn computational time of model is also significantly increased. Hence, there should be an appropriate mechanism for learning of training data. Such mechanism also helps to achieve contrastive divergence. It is also noticed that the training process become unstable due to higher learning rate, slightly lesser learning rate results in higher training time and slower convergence rate [27]. To deal with such issues, this work also presents an adaptive learning strategy for computing the learning rate. This strategy compute local learning rate (learning rate for each connection) instead of a single learning rate throughout an epoch. The independent learning rate parameter is obtained for each weight connection instead of a global learning rate to achieve the satisfactory training speed. The computation of learning parameter is summarized into equations (3.15-3.16)

$$\mu\varphi^a_{i,j}, if((VI_iHI_j)_k - (VI_iHI_j)_M)((VI_iHI_j)^a_K - ((VI_iHI_j))^{a-1}_m) > 0 \tag{3.15}$$

$$\vartheta\varphi^a_{i,j}, if((VI_iHI_j)_k - (VI_iHI_j)_M)((VI_iHI_j)^a_K - ((VI_iHI_j))^{a-1}_m) < 0 \tag{3.16}$$

The adaptive learning strategy for visible layer (VI) and hidden layer (HI) is illustrated in algorithm 1 and algorithm 2. In equation (3.15-3.16), $\mu > 0$ and $\vartheta < 0$ correspond to the increment and decrement factors of learning rate. The learning rate will be increased with two consecutive updates in similar direction; otherwise, it will be decreased. So, a uniform step size is introduced in the training process, in turn convergence speed can be improved. Further, the objective function for deep belief nets is described as error rate(ER(E)) and it is mentioned in equation (3.17).

$$ER(E) = \frac{1}{2}\sum_{i=1}^{k}\sum_{j=1}^{m}(\hat{y}_{i,j} - y_{i,j})^2 \tag{3.17}$$

Further, the process of training phase of deep belief nets with adaptive learning strategy is mentioned in Algorithm 3, while Algorithm 4 summarizes the testing phase process of deep belief nets.

3.3. Cloud Environment. The cloud environment is responsible for storing the data related to DR and communication among doctors and patients. It consists of large storage space to store data and this data can easily access at any time. The patients and family members can access health data of individual and also provide the feedback regarding the treatment. Other side, doctors access the patient data for treatment. A message facility is also enable in this environment regarding the health checkup, appointment and medical test.

Algorithm 1 Adaptive Learning Strategy for visible layer (VI)

for each visible layer (VI), i=1 to k **do**
  for each hidden layer (HI), j=1 to m **do**
    Compute the value of $\mu\phi_{(i,j)}{}^{a}$ using equation 3.15.
    if $(\mu > 0)$ **then**
      Learning rate $\phi_{(i,j)}{}^{a}$ is increased using previous value for two consecutive steps.
    end if
  end for
end for

Algorithm 2 Adaptive Learning Strategy for hidden layer (VI)

for each visible layer (VI), i=1 to k **do**
  for each hidden layer (HI), j=1 to m **do**
    Compute the value of $\mu\phi_{(i,j)}{}^{a}$ using equation 3.16.
    if $(\vartheta > 0)$ **then**
      Learning rate $\phi_{(i,j)}{}^{a}$ is decreased using previous value for two consecutive steps.
    end if
  end for
end for

Algorithm 3 Deep Belief Network with Adaptive Learning Strategy // Training Phase

Input: Training image dataset, number of epoch, number of RBM, number of visible layer, number of hidden layer.

Output: Diabetic Retinopathy Detection either presence or absence

Initialize the bias of visible and hidden layers (bs_i and bi_j), weight matrix($w_(p \times q)$), learning rate (φ), number of training data (N) and number of epoch (Max_Epoch).
while $(currentepoch \leq Max\_Epoch)$ **do**
  for each training data (c)=1 to N **do**
    for each visible layer (VI) i=1 to k **do**
      Compute the probability of hidden layer (HI) using equation 3.8.
    end for
    for each hidden layer (HI) j= 1 to m **do**
      Compute the probability of visible layer (HI) using equation 3.9.
    end for
    Assessed the hidden and visible layers probability using equation 3.10-3.11.
    Invoke the adaptive learning strategy (Algorithm 1)
    Update the parameters of RBM $(bi_j, bs_i, andw_{(i,j)})$ using equations 3.12-3.14.
  end for
  Training phase of RBM is completed and obtained the optimum tuning of parameters (bi_j, bs_i, and $w_{(i,j)}$) and learning rate (ϕ).
  Evaluate the objective function using equation 3.17 on output layer of deep belief nets.
  Back propagation method is adopted for computing diabetes in forward direction and weight in backward direction.
end while
Training of beep belief nets is completed and obtained the training accuracy.

4. Experimental Results. The section presents the effectiveness of the proposed diabetic retinopathy detection system. The efficacy of the proposed detection system is tested over retinal fundus images. In this, work, a total of three thousand two hundred retinal images are utilized for evaluating the performance of the

Algorithm 4 Deep Belief Network with Adaptive Learning Strategy // Testing Phase

Input: Test image dataset, number of epoch, number of RBM, number of visible layer, number of hidden layer.
Output: Diabetic Retinopathy Detection either presence or absence

Initialize the bias of visible and hidden layers (bs_i and bi_j), weight matrix($w_(p \times q)$), learning rate (φ), number of training data (N) and number of epoch (Max_Epoch).
while ($currentepoch \leq Max\_Epoch$) **do**
  for each testing data (c)=1 to N **do**
    for each visible layer (VI) i=1 to k **do**
      Compute the probability of hidden layer (HI) using equation 3.8.
    end for
    for each hidden layer (HI) j= 1 to m **do**
      Compute the probability of visible layer (HI) using equation 3.9.
    end for
    Assessed the hidden and visible layers probability using equation 3.10-3.11.
  end for
  Evaluate the objective function using equation 3.17 on output layer of deep belief nets.
  Back propagation method is adopted for computing diabetes in forward direction and weight in backward direction.
end while
Training of beep belief nets is completed and obtained the training accuracy.

proposed detection system in which two hundred twenty-four images having diabetic retinopathy symptoms and the rest of the images with no symptoms. The performance of the proposed diabetic retinopathy detection system is tested over a set of performance parameters. These parameters are accuracy, sensitivity, specificity, F1-Score and AUC. The training and validation accuracy along with loss rate are also investigated for over fitting issue of data. The abovementioned parameters are computed using the confusion matrix. The confusion matrix can be described in terms of true positive, true negative, false positive, and false negative. The proposed system is implemented in MATLAB 2016b environment using corei7 processor with 16GB RAM on window 10 operating system.

4.1. Results and Discussion. This subsection discusses the simulation results of the proposed model. Fig 4.1 shows the confusion matrix of the proposed diabetic retinopathy detection system and other techniques like InceptionV3, VGG19, VGG16, SVM, ANN and KNN. The confusion matrix is utilized to compute the other performance parameters such as accuracy, sensitivity, specificity, F1-score and AUC. Table 4.1 illustrate the experimental results of proposed diabetic retinopathy detection system and other techniques using accuracy and F1-score parameters.Further, these results are evaluated using Training/Testing (70%-30%), Training/Testing (80%-20%), 5-CrossFold Validation, and 10-CrossFold Validation methods. In Training/Testing (70%-30%), the entire data is divided into 70% and 30%. The seventy percent data is considered for training set and adopted for train the model, while thirty percent dataset is used as validation set and applied for evaluating the performance of the model.

In Training/Testing (80%-20%), eighty percent of entire data is adopted for training set, while twenty percent data is considered for validation set and can be used to evaluate the performance of model. In 5-cross fold validation, the entire dataset is divided into five equal sized set and out of five, four sets are applied to train the model as training set, rest one is used to evaluate the performance of model as validation set. This process is repeated up to five times, but every time validation set is different. In 10-cross fold validation, the data is divided into ten equal sized sets in which nine sets are used as training set and tenth set is employed for evaluating the performance of model as validation set. This process is repeated up to ten times and every times validation set is different. It is found that proposed diabetic retinopathy detection system achieves better results than other techniques like InceptionV3, VGG19, VGG16, SVM, ANN and KNN in terms of accuracy and F1- Score parameters using all aforementioned methods. The accuracy and F1-score rates of proposed diabetic retinopathy detection system using Training/Testing (70%-30%), Training/Testing (80%-20%), 5-CrossFold

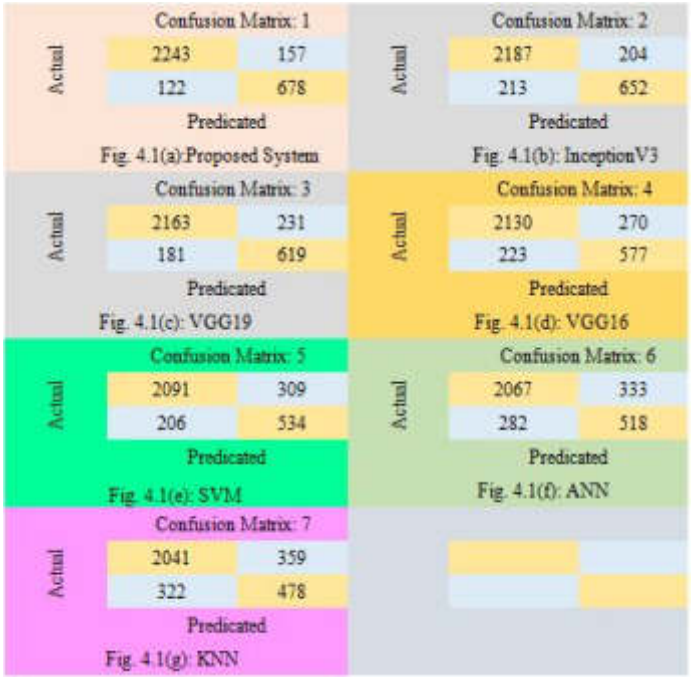


Fig. 4.1: (a-g) Depicts the Confusion Matrix of proposed diabetic retinopathy detection system and other techniques.

Table 4.1: Results of proposed model and other techniques for diabetic retinopathy using accuracy and F1-score parameters.

Technique	Training/Testing (70%-30%)		Training/Testing (80%-20%)		5- Cross Fold Validation		10-Cross Fold Validation	
	Accuracy	F1-Score	Accuracy	F1-Score	Accuracy	F1-Score	Accuracy	F1-Score
KNN	69.21	71.23	73.21	76.06	76.46	82.95	78.72	85.72
ANN	70.93	72.06	75.64	79.26	77.09	84.88	80.78	87.04
SVM	73.24	76.76	77.52	80.03	80.41	85.73	82.03	87.91
VGG16	74.67	77.18	78.97	81.12	82.94	87.81	84.59	89.63
VGG19	77.03	78.73	81.53	83.05	85.91	89.56	86.94	91.31
Inception V3	79.89	81.78	84.93	86.53	87.09	90.14	88.72	91.38
Proposed System	85.91	87.48	87.41	89.11	90.46	92.5	91.28	94.14

Validation, and 10-CrossFold Validation methods are (85.91 and 87.48), (87.41 and 89.11), (90.46 and 92.5), (91.28 and 94.14) respectively. It is analyzed that proposed system having higher accuracy and F1-score rate than other techniques using all possible training/testing and cross fold validation methods. It is also observed that KNN technique provides less efficient results for detecting of the diabetic retinopathy in terms of accuracy and F1-score as (69.21 and 71.23) with training/testing (70%-30%), (73.21 and 76.06) with training/testing (80%-20%), (76.46 and 82.95) with 5-cross fold validation, and (78.22 and 85.72) with 10-cross fold validation method. It is also revealed that ANN and SVM techniques obtain similar F1-score rates using 10-cross fold validation method. Hence, it stated that proposed diabetic retinopathy detection system is provide more accurate results in terms of accuracy and F1-score rate for detection of diabetes retinopathy. It is also seen that proposed detection model achieves more than 94% F1-score rate using 10-cross fold validation method for diabetic retinopathy. Further, 10-cross fold validation method significantly enhances the accuracy and F1-score

Table 4.2: Results of proposed model and other techniques for diabetic retinopathy using sensitivity and specificity parameters.

Technique	Training/Testing (70%-30%)		Training/Testing (80%-20%)		5- Cross Fold Validation		10-Cross Fold Validation	
	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity
KNN	71.42	71.04	74.28	77.92	82.61	83.29	85.04	86.37
ANN	71.04	73.12	78.34	80.21	84.73	85.03	86.12	87.99
SVM	75.34	78.23	78.73	81.37	85.34	86.12	87.13	88.71
VGG16	76.81	77.56	80.03	82.23	87.24	88.38	88.75	90.52
VGG19	77.43	80.07	82.64	83.46	88.41	90.74	90.35	92.28
Inception V3	80.21	83.41	86.02	87.04	89.83	90.45	91.47	91.13
Proposed System	86.73	88.24	88.13	90.11	92.04	92.96	93.46	94.84

results more than 6% and 7% in comparison to training/testing (70%-30%) method. Sensitivity and specificity parameters are also considered for evaluating the performance of the proposed diabetic retinopathy detection system. These parameters are evaluated the performance of the model in term of true positive rate i.e. actual presence of the diabetic retinopathy with respect to false positive and false negative. The experimental results of proposed diabetic retinopathy detection system and other techniques based on sensitivity and specificity are reported into Table 4.2.

It is observed that proposed diabetic retinopathy detection system achieves higher sensitivity and specificity rate such as (86.73 and 88.24) with training/testing (70%-30%) method, (88.13 and 90.11) with training/testing (80%-20%) method, (92.04 and 92.96) with 5-cross fold validation method, and (93.46 and 94.84) with 10-cross fold validation method. Similar, KNN technique obtains less accurate results in terms of sensitivity and specificity for diabetic retinopathy using training/testing and cross validation methods. It is seen that the sensitivity and specificity rates of KNN technique with Training/Testing (70%-30%), Training/Testing (80%-20%), 5-CrossFold Validation, and 10-CrossFold Validation are (71.42 and 71.04), (74.28 and 77.92), (82.61 and 83.29), and (85.04 and 86.37) respectively. It is also seen that ANN technique provides minimum sensitivity rate (71.04) with training/testing (70%-30%) method as compared to other techniques. It is also stated that 10-cross fold validation is more significant method as compared to other methods like Training/Testing (70%-30%), Training/Testing (80%-20%),and 5-CrossFold Validation methods as this method improve the sensitivity and specificity rates up to more than 6% and 7% in comparison to training/testing (70%-30%) method. The accuracy, F1-score, sensitivity and specificity rates based on training/testing and validation methods using all techniques are presented into Fig 4.2-4.5. The accuracy results of proposed diabetic retinopathy detection system and other techniques are showed into Fig 4.2.

It is revealed that proposed detection system obtains higher accuracy results using all training/ testing and validation methods in comparison to other techniques.It is also analyzed that among all training/testing and validation methods, the training/testing (70%-30%) method with all techniques including proposed diabetic retinopathy detection system provides less accurate results than 10-cross fold validation method. It is also stated that 10-cross validation method significantly improves the experimental results of all techniques.

Fig 4.3 presents the F1-score results of proposed diabetic retinopathy detection system and other techniques using all methods. Similarly, 10-cross validation method having better F1-score rate than other all techniques. The experimental results of sensitivity and specificity parameter using training/testing and validation methods are demonstrated into Fig 4.4-4.5. It is revealed that 10-cross fold validation method significantly improves the simulation results of all technique as compared training/testing (70%-30%), training/testing (80%-20%), and 5-cross fold validation methods. It is also found that all techniques exhibit less accurate results with training/testing (70%-30%) method. Hence, it is concluded that 10-cross validation method is an effective method for assessing the simulation results of techniques for diabetic retinopathy. Furthermore, it is found that proposed diabetic retinopathy detection system obtains better results for diabetic retinopathy with all training/testing and validation method as compared to other techniques.

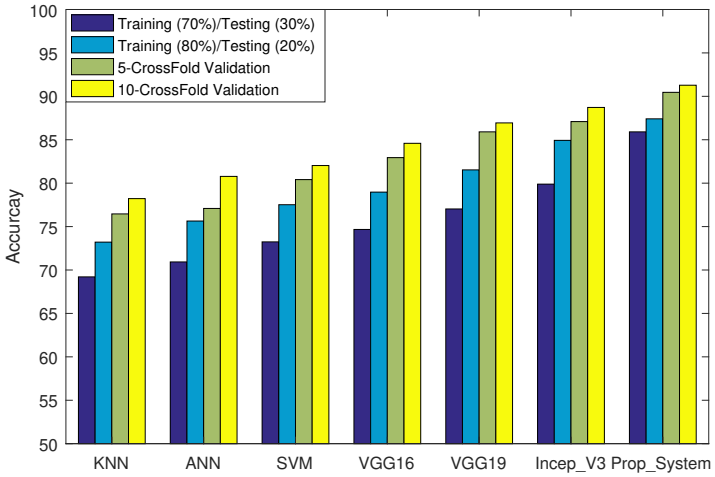


Fig. 4.2: Illustrates the accuracy results of proposed diabetic retinopathy detection system and other techniques using training /testing and validation methods.

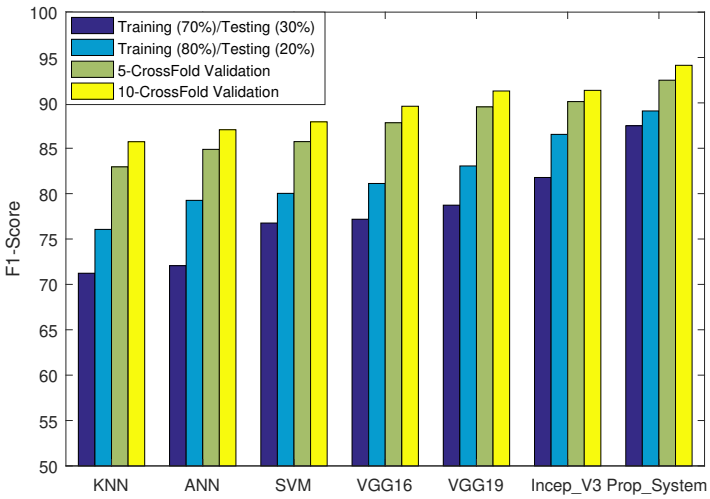


Fig. 4.3: Illustrates the F1-score results of proposed diabetic retinopathy detection system and other techniques using training /testing and validation methods.

Further, the experimental results of AUC parameter for proposed diabetic retinopathy detection system and other technique are presented into Fig 4.6. The AUC result of proposed diabetic retinopathy detection system is compared with KNN, SVM, ANN, VGG16, VGG19 and InceptionV3 techniques. The AUC parameter is defined in terms of true positive rate (TPR) and false positive rate (FPR). The TPR and FPR are used to plot the AUC result which is shown in Fig 4.6. It is stated that proposed model achieves higher AUC rate than other techniques. Hence, this parameter showed the effectiveness of the proposed diabetic retinopathy detection system for detecting the diabetic retinopathy. Finally, it is stated that proposed diabetic retinopathy detection system achieves more accurate results in terms of accuracy, sensitivity, specificity, F1-score and AUC parameters than other techniques.

5. Conclusion. This paper presents a diabetic retinopathy detection system for accurate diagnosis of diabetes retinopathy through retinal fundus images. The efficacy of the proposed diabetic retinopathy detection

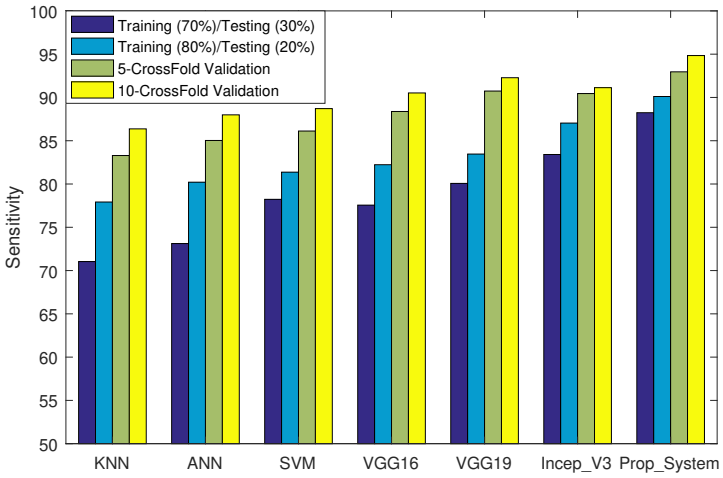


Fig. 4.4: Illustrates the sensitivity results of proposed diabetic retinopathy detection system and other techniques using training /testing and validation methods.

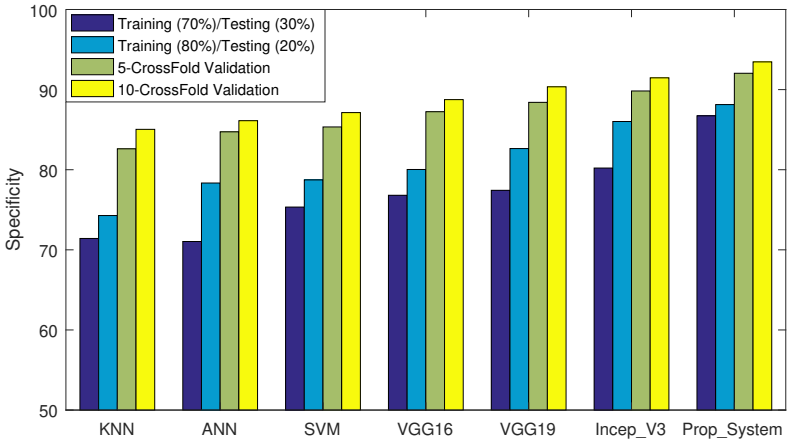


Fig. 4.5: Illustrates the specificity results of proposed diabetic retinopathy detection system and other techniques using training /testing and validation methods.

system is evaluated on three thousand two hundred retinal fundus images in which twenty four images are affected with diabetic retinopathy and rest of images are without diabetes retinopathy. Several preprocessing techniques are employed on image dataset in terms of threshold and luminosity, grayscale, edge detection, histogram and k-means based segmentation technique. In turn, processed dataset is constructed for detecting the diabetes retinopathy. In this work, deep belief nets are adopted for accurate diagnosis of diabetes retinopathy. Prior to implement the deep belief nets, an adaptive learning strategy is integrated into deep belief nets for computing the optimal learning rate in each iteration. The experimental results are evaluated using accuracy, sensitivity, specificity, F1-score and AUC parameters and also compared with state of art existing techniques like ANN, KNN, SVM, VGG16, VGG19, InceptionV3. Moreover, the experimental results of proposed are also assessed using Training/Testing (70%-30%), Training/Testing (80%-20%), 5-CrossFold Validation, and 10-CrossFold Validation methods. It is found that proposed diabetic retinopathy detection system achieves more accurate results than other techniques in terms of accuracy, sensitivity, specificity, F1-score and AUC parameter. It is also notice that proposed system obtains 94.14% F1-score rate using 10-cross fold validation method. On the

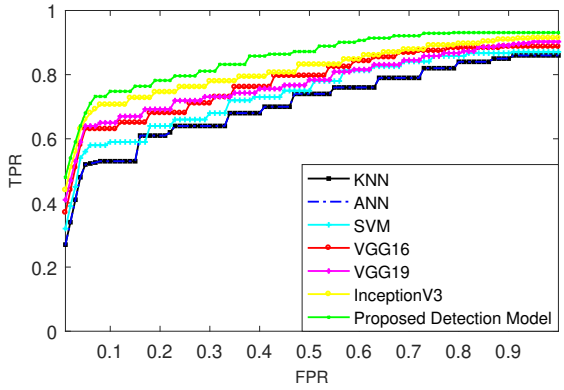


Fig. 4.6: Illustrates the area under curve (AUC) results of proposed diabetic retinopathy detection system and other techniques.

analysis of different training/testing and validation methods, it is found that 10-cross fold validation method achieve better accurate results than other methods, while training/testing (70%-30%) method having less accurate results for detection of diabetic retinopathy. Finally, it is concluded that proposed diabetic retinopathy detection system provides better results than other techniques using all training/testing and validation methods. Hence, it is stated that proposed diabetic retinopathy detection system is an effective and efficient technique for detecting diabetes retinopathy.In future, meta-heuristic algorithm-based segmentation techniques will be explored to determine the region of interest. This work considers the statistical and spatial features to design the diabetic retinopathy dataset. In future work, the features from fundus images will be extracted using first-order and second-order derivatives. Recently developed meta-heuristic algorithms can utilize to detect diabetic retinopathy more accurately and effectively.

REFERENCES

[1] Q. Abbas, I. Fondon, A. Sarmiento, S. Jiménez, and P. Alemany, *Automatic recognition of severity level for diagnosis of diabetic retinopathy using deep visual features*, Medical & biological engineering & computing, 55 (2017), pp. 1959–1974.

[2] W. L. Alyoubi, M. F. Abulkhair, and W. M. Shalash, *Diabetic retinopathy fundus image classification and lesions localization system using deep learning*, Sensors, 21 (2021), p. 3704.

[3] Y. Bengio, P. Lamblin, D. Popovici, and H. Larochelle, *Greedy layer-wise training of deep networks*, Advances in neural information processing systems, 19 (2006).

[4] M. Canayaz, *Classification of diabetic retinopathy with feature selection over deep features using nature-inspired wrapper methods*, Applied Soft Computing, 128 (2022), p. 109462.

[5] M. A. Carreira-Perpinan and G. Hinton, *On contrastive divergence learning*, in International workshop on artificial intelligence and statistics, PMLR, 2005, pp. 33–40.

[6] J. J. G. Chandran, J. Jabez, and S. Srinivasulu, *Auto-metric graph neural network optimized with capuchin search optimization algorithm for coinciding diabetic retinopathy and diabetic macular edema grading*, Biomedical Signal Processing and Control, 80 (2023), p. 104386.

[7] Ö. Deperlioğlu and U. Köse, *Diagnosis of diabetic retinopathy by using image processing and convolutional neural network*, in 2018 2nd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), IEEE, 2018, pp. 1–5.

[8] Z. Gao, J. Li, J. Guo, Y. Chen, Z. Yi, and J. Zhong, *Diagnosis of diabetic retinopathy using deep neural networks*, IEEE Access, 7 (2018), pp. 3360–3370.

[9] R. Gargeya and T. Leng, *Automated identification of diabetic retinopathy using deep learning*, Ophthalmology, 124 (2017), pp. 962–969.

[10] N. Gundluru, D. S. Rajput, K. Lakshmanna, R. Kaluri, M. Shorfuzzaman, M. Uddin, and M. A. Rahman Khan, *Enhancement of detection of diabetic retinopathy using harris hawks optimization with deep learning model*, Computational Intelligence and Neuroscience, 2022 (2022).

[11] V. M. G. S. Gupta, S. Gupta, P. Sengar, et al., *Extraction of blood veins from the fundus image to detect diabetic retinopathy*, in 2016 IEEE 1st International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES), IEEE, 2016, pp. 1–3.

[12] Ö. F. GÜRCAN, U. ATICI, AND Ö. F. BEYCA, *A hybrid deep learning-metaheuristic model for diagnosis of diabetic retinopathy*, Gazi University Journal of Science, 36, pp. 693–703.
[13] D. J. HEMANTH, J. ANITHA, L. H. SON, AND M. MITTAL, *Diabetic retinopathy diagnosis from retinal images using modified hopfield neural network*, Journal of medical systems, 42 (2018), pp. 1–6.
[14] D. J. HEMANTH, O. DEPERLIOGLU, AND U. KOSE, *An enhanced diabetic retinopathy detection and classification approach using deep convolutional neural network*, Neural Computing and Applications, 32 (2020), pp. 707–721.
[15] G. E. HINTON, *Deep belief networks*, Scholarpedia, 4 (2009), p. 5947.
[16] T. J. JEBASEELI, C. A. D. DURAI, AND J. D. PETER, *Iot based sustainable diabetic retinopathy diagnosis system*, Sustainable Computing: Informatics and Systems, 28 (2020), p. 100272.
[17] J. P. KANDHASAMY, S. BALAMURALI, S. KADRY, AND L. K. RAMASAMY, *Diagnosis of diabetic retinopathy using multi level set segmentation algorithm with feature extraction using svm with selective features*, Multimedia Tools and Applications, 79 (2020), pp. 10581–10596.
[18] H. KAUSHIK, D. SINGH, M. KAUR, H. ALSHAZLY, A. ZAGUIA, AND H. HAMAM, *Diabetic retinopathy diagnosis from fundus images using stacked generalization of deep models*, IEEE Access, 9 (2021), pp. 108276–108292.
[19] R. F. MANSOUR, *Deep-learning-based automatic computer-aided diagnosis system for diabetic retinopathy*, Biomedical engineering letters, 8 (2018), pp. 41–57.
[20] P. MODI AND Y. KUMAR, *Smart detection and diagnosis of diabetic retinopathy using bat based feature selection algorithm and deep forest technique*, Computers & Industrial Engineering, (2023), p. 109364.
[21] L. QIAO, Y. ZHU, AND H. ZHOU, *Diabetic retinopathy detection using prognosis of microaneurysm and early diagnosis system for non-proliferative diabetic retinopathy based on deep learning algorithms*, IEEE Access, 8 (2020), pp. 104292–104302.
[22] I. QURESHI, J. MA, AND Q. ABBAS, *Diabetic retinopathy detection and stage classification in eye fundus images using active deep learning*, Multimedia Tools and Applications, 80 (2021), pp. 11691–11721.
[23] M. RAGAB, W. H. ALJEDAIBI, A. F. NAHHAS, AND I. R. ALZAHRANI, *Computer aided diagnosis of diabetic retinopathy grading using spiking neural network*, Computers and Electrical Engineering, 101 (2022), p. 108014.
[24] A. SHARAFELDEEN, M. ELSHARKAWY, F. KHALIFA, A. SOLIMAN, M. GHAZAL, M. ALHALABI, M. YAGHI, M. ALRAHMAWY, S. ELMOUGY, H. SANDHU, ET AL., *Precise higher-order reflectivity and morphology models for early diagnosis of diabetic retinopathy using oct images*, Scientific Reports, 11 (2021), pp. 1–16.
[25] K. SIMONYAN AND A. ZISSERMAN, *Very deep convolutional networks for large-scale image recognition*, arXiv preprint arXiv:1409.1556, (2014).
[26] A. SKOUTA, A. ELMOUFIDI, S. JAI-ANDALOUSSI, AND O. OUCHETTO, *Hemorrhage semantic segmentation in fundus images for the diagnosis of diabetic retinopathy by using a convolutional neural network*, Journal of Big Data, 9 (2022), pp. 1–24.
[27] P. SOKKHEY AND T. OKAZAKI, *Development and optimization of deep belief networks applied for academic performance prediction with larger datasets*, IEIE Transactions on Smart Processing & Computing, 9 (2020), pp. 298–311.
[28] D. S. W. TING, G. C. M. CHEUNG, AND T. Y. WONG, *Diabetic retinopathy: global prevalence, major risk factors, screening practices and public health challenges: a review*, Clinical & experimental ophthalmology, 44 (2016), pp. 260–277.
[29] T. VIJAYAN, M. SANGEETHA, A. KUMARAVEL, AND B. KARTHIK, *Feature selection for simple color histogram filter based on retinal fundus images for diabetic retinopathy recognition*, IETE Journal of Research, (2020), pp. 1–8.
[30] W. A. WAN MUSTAFA, H. YAZID, AND M. M. M. ABDUL KADER, *Luminosity correction using statistical features on retinal images*, in Journal of Biomimetics, Biomaterials and Biomedical Engineering, vol. 37, Trans Tech Publ, 2018, pp. 74–84.
[31] J. WANG, Y. BAI, AND B. XIA, *Simultaneous diagnosis of severity and features of diabetic retinopathy in fundus photography using deep learning*, IEEE Journal of Biomedical and Health Informatics, 24 (2020), pp. 3397–3407.
[32] X. WANG, M. XU, J. ZHANG, L. JIANG, L. LI, M. HE, N. WANG, H. LIU, AND Z. WANG, *Joint learning of multi-level tasks for diabetic retinopathy grading on low-resolution fundus images*, IEEE Journal of Biomedical and Health Informatics, 26 (2021), pp. 2216–2227.
[33] R. E. WASHINGTON, T. J. ORCHARD, V. C. ARENA, R. E. LAPORTE, A. M. SECREST, AND E. S. TULL, *All-cause mortality in a population-based type 1 diabetes cohort in the us virgin islands*, Diabetes research and clinical practice, 103 (2014), pp. 504–509.
[34] G. YUE, Y. LI, T. ZHOU, X. ZHOU, Y. LIU, AND T. WANG, *Attention-driven cascaded network for diabetic retinopathy grading from fundus images*, Biomedical Signal Processing and Control, 80 (2023), p. 104370.
[35] G. T. ZAGO, R. V. ANDREÃO, B. DORIZZI, AND E. O. T. SALLES, *Diabetic retinopathy detection using red lesion localization and convolutional neural networks*, Computers in biology and medicine, 116 (2020), p. 103537.
[36] H. ZHANG, Q. M. J. WU, Y. ZHENG, T. M. NGUYEN, AND D. WANG, *Effective fuzzy clustering algorithm with bayesian model and mean template for image segmentation*, IET Image Processing, 8 (2014), pp. 571–581.
[37] L. ZHOU, Y. ZHAO, J. YANG, Q. YU, AND X. XU, *Deep multiple instance learning for automatic detection of diabetic retinopathy in retinal images*, IET Image Processing, 12 (2018), pp. 563–571.
[38] M. ZHOU, K. JIN, S. WANG, J. YE, AND D. QIAN, *Color retinal image enhancement based on luminosity and contrast adjustment*, IEEE Transactions on Biomedical engineering, 65 (2017), pp. 521–527.

Edited by: Pradeep Kumar Singh
Received: Feb 25, 2023
Accepted: Jul 28, 2023


Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2023 SCPE. Volume 24, Issues 2, pp. 115–125, DOI 10.12694/scpe.v24i2.2145

INFORMATION MONITORING OF TRANSMISSION LINES BASED ON INTERNET OF THINGS TECHNOLOGY

WEI DU, JIAYING WANG, GUOZHU YANG, SIJIA ZHENG, AND YAJIE ZHAO*

Abstract. To solve the problem of difficult real-time monitoring of current transmission lines, this article proposes an information-based monitoring system for transmission lines based on Internet of Things technology. The system utilizes the characteristics of strong scalability, good fault tolerance, low power consumption, and low cost of the Internet of Things. Taking the ultra-low power consumption MSP430 microcontroller and CC2430 radio frequency module as the core, a line monitoring system based on the Internet of Things is designed. The proposed design uses ZigBee wireless sensor network technology which is powered by solar energy. The collection, transmission, processing and judgment of various environmental parameters of the line are realized. The data information is transferred to the monitoring center of the upper computer through GPRS. When there is an abnormality, it can send a mobile phone short message to the person in charge to feedback the abnormal content in time. The distribution network's load symmetry allowed for the development of several locating procedures. For the three-phase symmetric scheme, the fault location approach based on line supply characteristics was employed, and for the three-phase asymmetric scheme, the fault location technique based on line impedance is proposed. One of the most vital uses for the Internet of Things is in the mitigation of power transmission line failures and disasters. Improved power transmission dependability, less financial loss, and fewer power outages are all possible thanks to the Internet of Things' cutting-edge sensing and communication technology. This research introduced the use of IoT in online monitoring system of electricity transmission line with a focus on the characteristics of the construction and development of smart grid. The results indicated that the system's highest temperature difference is 0.31°C, while the maximum humidity difference is 1.38%. The system increases the safety and manageability of electricity transmission while also fostering the widespread adoption and technical integration of the smart grid and the Internet of Things.

Key words: Internet of things; CC2430; Smart grid; GPRS/GSM; Data acquisition, Internet of Things (IOT)

1. Introduction. The Internet of Things technology is a technology that has emerged with the trend of technological development and people's expectations for high living standards, this technology can better ensure people's electricity safety and monitor the power grid at any time. To ensure people's requirements for a high standard of living. And the detection system under the Internet of Things technology can detect various problems at any time. Therefore, it is very necessary to use the Internet of Things technology to monitor the transmission line online. Also, due to the needs of the application, it is necessary to make a brief introduction to Internet technology. To better understand the meaning and role of the Internet of Things among users and beneficiaries. The economic development of the country is inextricably linked with the support of the electricity industry [1]. As technology advances in the world, the latest Internet technologies are gradually being introduced into electric transportation applications. Based on the Internet of Things technology, the online monitoring of the transmission line can detect the detailed problems of the transmission line anytime, anywhere, and reduce unnecessary losses to the maximum extent. Therefore, the online monitoring system of transmission lines using the Internet of Things technology is a necessary trend for national development and reform and world progress. Intelligent technology is the mainstream of technological development in the 21st century. Using various intelligent technologies to solve a series of problems in life and production can effectively reduce manual labor and save time. And in the overall inspection management, has a role that cannot be ignored.

The application of intelligent technology to transmission lines is a bold reform and innovation. The detection system can effectively prevent some problems that are easily overlooked [2]. The detection system of the Internet of Things can be updated at any time to check the transmission line and reduce the occurrence of accidents.

*State Grid Electric Power Space Technology Company Limited, Beiiing,102209, China (weidu1824@gmail.com, wjiaying58@gmail.com, guozhuyang99@gmail.com, sijiazheng7@gmail.com, zhaoyajie750@gmail.com). Corresponding author: Sijia Zheng

Therefore, the application of the online monitoring system of the transmission line based on the Internet of Things technology is imperative. The Internet of Things, as the name suggests, is a network of connected objects. But in fact, it is not just the connection between objects and objects, but also the connection between people and objects. But no matter what kind of connection method, you need to intervene in the wireless network to play its due role. And connect the network in time for the relevant signs of the objects. Key features of the Internet of Things. The Internet of Things is a type of control and monitoring system that connects networks and provides intelligence to objects. This mode can identify various transmission problems more effectively and solve problems in time [3]. Smart grid is a hardware product because it is the new technological development of the new century. Therefore, understanding the importance and nature of the Internet of Things, that is, the nature of the information connection between things and people and things, can be useful in your work. Comprehensive state awareness, effective information processing, and practical and adaptable application are the traits of an IoT. It also contributes to the manufacture, consumption, operation, facility, and other features of energy in order to satisfy the strains of the supply network's multitype application. The fault segment location approach and the fault precise location method are currently the foremost themes of study on power system fault location. A D-PMU is the PMU application in the supply system and serves as an essential measurement expedient in the IoT. The voltage and current indications of the supply system can be calculated by a D-PMU in both steady state and abnormal circumstances. Numerous approaches for locating faults have been put forth, counting methods based on neural networks, upgraded line impedance methods, and methods based on travelling waves.

1.1. Problem Statement and Motivation. The data information is transferred to the monitoring center of the upper computer through GPRS. When there is an abnormality, it can send a mobile phone short message to the person in charge to feedback the abnormal content in time. The distribution network's load symmetry allowed for the development of several locating procedures. Overhead high-voltage transmission lines are susceptible to wind vibration and wind deviation, which are the leading causes of transmission line failure. Strong winds can generate conductor galloping, which can occur for many hours and inflict significant damage to high-voltage transmission lines [4]. Possible threats to transmission line safety include ice due to rain and snow, tilting of the transmission tower due to unequal pulling forces, and power outages caused by these factors. The weather conditions along the transmission corridor are not reflected in the monitoring record supplied by the local meteorological station at any given moment. In addition, there are obstacles to fault diagnosis, prevention, and study of the power transmission line since there is essentially no historical weather data for the transmission corridor [5]. Internet of Things (IoT) technology used for online monitoring of power transmission lines provides the key to addressing the aforementioned issues.

1.2. Contribution and Organization. Transmission lines are an integral aspect of the electricity system, and their efficiency is a major factor in how smoothly the grid functions overall. The monitoring of the status information of power equipment based on the Internet of Things has significant research value because of the ongoing promotion of the application of the Internet of Things technology in the power industry. In this article, we combine theory and practise to examine the data information, tasks, and technical methods of power equipment status monitoring; we then conduct in-depth studies of both primary and secondary machines, and we conclude by outlining a strategy and method for monitoring power equipment status information via the Internet of Things [6]. The remaining article is structured as follows: a literature review is presented in Section 2 of the article, followed by a discussion of research methods, an overview of smart grid and IoT, the overall design scheme of the system, system hardware design, and system software design are explained in Section 3. Section 4 presents the results and discussion, followed by the conclusion section in Section 5.

2. Literature Review. In recent years, with the development of the economy, the country's electricity consumption has increased rapidly, and the construction and measurement of ultra-high voltage power system links have been rapidly expanding. Due to the characteristics of long distance, wide distribution, and difficult inspection and maintenance of high-voltage transmission lines, the transmission line network operating in remote areas with complex terrain and harsh environments, real-time remote monitoring of transmission lines and their environmental meteorological parameters has become an urgent task, by monitoring the operation status of transmission lines in real time and establishing a corresponding natural disaster early warning mechanism, the

economic losses caused by power outages can be reduced, and the safety, stability, and efficiency of power grid operation can be improved [7]. Therefore, the establishment of online monitoring of transmission lines has played an important role in the stable operation of the power grid, and online monitoring of transmission lines is an important part of the smart grid. The traditional monitoring network is mainly wired, and there are problems such as complicated wiring, low reliability, low security, high cost, and difficulty in expansion and maintenance [8]. With the existing commonly used wireless communication technologies such as GPRS, and WiFi, there are problems such as high cost and high-power consumption.

With the goals of low power consumption, cheap cost, security, reliability, high network capacity, high performance, wireless, easy expansion, and electrical transmission in mind, Yang et al. [9] suggested a design based on wireless sensor network and ZigBee technology. Its ease of usage has led to its growing incorporation into the infrastructure controlling transmission lines. ZigBee networks may be integrated with preexisting communication infrastructure, and the usual 80-meter distance between network nodes can be increased to hundreds of metres or even kilometres with the use of power amplifiers [10]. In China, infection control systems are evolving quickly with advances in sensor, data transmission, and AI technology. The system uses Internet of Things technology to set up a multi-sensor cooperative wireless sensor network by attaching smart sensors to towers, transmission lines, and other crucial electronic equipment.

Parameters such as wire temperature, humidity, sag, line icing, breeze vibration, wire wind deflection, tower inclination, insulator contamination, etc. are collected through sensors, the data is transmitted to the upper computer monitoring center by the combination of optical fiber communication and wireless communication.When the parameters are abnormal, it will automatically alarm, reminding the duty personnel and management personnel to take relevant measures in time to avoid accidents [11]. Considering the requirements of the working environment and conditions of high-voltage transmission line monitoring equipment, the author designed a wireless transmission line online monitoring system based on ZigBee technology [12]. Combined with the inductive method, it easily solves the problem of high-voltage insulation, and at the same time, it has the advantages of a large capacity network, small size, lightweight, low energy consumption, and easy installation. good. the ideal solution for controlling high-voltage transmission.A fault location approach based on online impedance was published in [13], and it was applicable for both multi-source and single-source power distribution systems because it required less data. The approaches put forth by [14] and [15] have great accuracy, but they failed to take grid operation symmetry into account when determining fault location accuracy. The use of synchronized voltage and current phasors obtained from phasor measuring units or clever electronic devices as a novel fault-finding technique for multi-terminal nonhomogeneous transmission lines was reported in [16]. The Internet of Things (IoT) is an extensive network of interconnected computing and sensing nodes that collect and interchange data using technologies such as radio-frequency identification (RFID) tags, infrared sensors, GPS, laser scanners, and the Internet itself to collect data. The IoT collects data and verifies identities using a variety of sensors and smart devices that are interconnected. In order to analyze data and unearth concealed insights, it relies on the worldwide web and other communication networks and employs numerous computer systems and applications [17]. Through the exchange of data between humans and objects, as well as between objects themselves, IoT technology enables real-time control, precise administration, and scientific decision-making of the physical world. In the smart grid, Power IoT (PIoT) represents the Internet of Things. With the assistance of a wired or wireless communication network and intelligent data processing in the power grid system, PIoT can achieve its objective of dependable information transfer [18, 19]. Power generation, transmission, transformation, distribution, and consumption are only a few of the numerous use cases for the Internet of Things in a smart grid. When high-voltage transmission lines are exposed to the elements, transmission line damage, interference with the secure operation of transmission facilities, the inability to provide power to a large region, and a substantial loss to the national economy are all possible outcomes.

3. Research methods.

3.1. Overview of Smart Grid and IoT. The smart grid consists of integrated, high-speed two-way communication, and is a new modern energy technology developed by integrating advanced measurement technology, information communication technology, analysis and decision-making technology, and automatic control. technology and energy and energy technology [20]. The purpose of the smart grid is to ensure security, safety, economy, efficiency, environmental performance, and safety, and its main characteristics are robust-

ness, self-healing, socialization, financial efficiency, integration, and optimization.Compared with the existing energy grid, the smart grid integrates energy flow, information flow, and business flow very effectively, and its advantages are generally as follows:

1. It integrates and improves the power of the power grid, which is resilient to various external influences and attacks, adaptable to the use of clean energy and renewable energy, and has a strong infrastructure for power generation and support.
2. Information technology, sensor technology, automatic control technology, integration of electronic plans, collection and reception of electrical work characteristics, timely detection and prediction of faults. In the event of a fault, the power grid can quickly isolate the fault, restore itself, and avoid massive power outages [21].
3. Through the use of modern technology of communication, information, and management, it can improve the efficiency of the use of energy resources, reduce energy loss, and make the operation of the electric grid energy-saving and efficient.
4. Provide clear, complete, and state-of-the-art operational plans for advanced data integration, sharing, use, and performance management, including support decisions, management plans, and response plans.
5. Create a two-way interactive service model, users can understand information about energy capacity, energy quality, energy cost, and energy efficiency in timely, and effective energy management; Energy companies can obtain detailed information on consumers energy consumption and provide value-added services [22].

3.2. Overall design scheme of the system. The system is mainly composed of two parts: line monitoring system (lower computer) and a monitoring center (upper computer). The line monitoring system is mainly composed of various sensor nodes (terminal nodes), cameras, main processing units (coordinator nodes), and solar power supply units. Among them, the sensor nodes can be configured as required, usually including tension sensors, wind speed sensors, wind direction sensors, temperature and humidity sensors, inclination sensors, leakage current sensors, and wire temperature sensors. The main processing unit and the solar power supply unit are generally installed on the line tower, and the sensor node is installed at the position that needs to be monitored, the sensor node has its wireless module, which can wirelessly send the collected data to the main processing unit, after processing and analysis, the relevant data and images are transmitted to the upper computer monitoring center through the power line or GPRS network [23]. The host computer stores the current image data and related sensor data in the database or hard disk, and the user of the power monitoring network accesses the receiving center through the monitoring software, and reads various parameter information of the current tower and line from the database or hard disk, provide a basis for early warning of transmission line disasters. The system mainly includes the functions of icing state monitoring, meteorological parameter monitoring, image monitoring, electrical parameter monitoring, mechanical parameter monitoring, and wire temperature monitoring. Ice-covering state monitoring: Measure the wire weight, insulator declination, wind speed, wind direction, temperature, and other parameters through relevant online monitoring equipment, the analysis software synthesizes the above data and wire parameters using relevant mathematical models, the equivalent ice thickness of the wire is calculated [24]. Meteorological parameter monitoring: Micrometeorological monitoring online monitoring of parameters such as line ambient temperature, humidity, wind speed, wind direction, rainfall, atmospheric pressure, etc., comprehensively analyze meteorological data and display all data to users through various reports, statistical charts, curves, etc.

The coordinator processing unit and the solar power supply system are installed on the tower, and the terminal sensor unit is installed on the line, in which the coordinator node is the core part, of the system, it is responsible for establishing a network and communicating with each sensor node joining the network, receiving data, issuing control commands, and transmitting the collected data to the monitoring center through GPRS or power line; The terminal sensor is configured as required, and is mainly responsible for the collection and transmission of on-site environmental parameters. It consists of sensors, alarms, and ZigBee modules. Due to its small size, low power consumption, and strong function, it can be fixed on high-voltage transmission lines, and the built-in solar power supply device can meet the power demand. The sensor node collects various environmental and state parameters of the transmission line, then the process and volume of the data are stored and sent wirelessly to the partner. When transmitted to the control center of GPRS, the control center

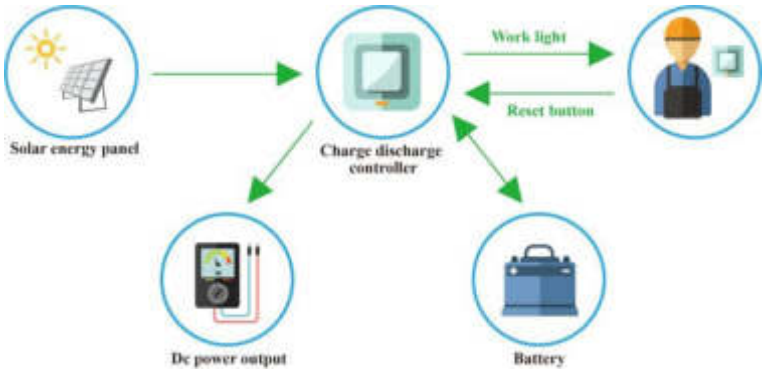


Fig. 3.1: Structure diagram of solar power supply

processes and collects data and reports it in real-time, so that when the transmission is abnormal, control of the timing of transmission lines can be used. will automatically alert lines, send reports to the control center and responders, and take timely measures to prevent power outages and other incidents layer [25].

3.3. System Hardware Design.

3.3.1. Power supply unit hardware. Due to the lack of an easily accessible AC power source for the high-voltage gearbox line in the wild, a battery-powered line monitoring device is impractical. There are now two major ways of power supply for high-voltage transmission line monitoring systems, thus it is important to design an independent power supply device with the aid of energy harvesting technology. The energy of the rotating electromagnetic field around the high-voltage AC wire may be harvested via the concept of electromagnetic induction, and then the alternating voltage produced can be transformed into direct current and either used immediately or stored in a battery. The second option is to use the solar energy system to provide power. Electromagnetic induction is challenging to regulate precisely because of its high initial current. As a result, solar panels and charge and discharge controllers make up the system's solar power source, as seen in Figure 3.1. The charge and discharge controller takes the solar panel's output voltage and converts it to the stable DC voltage needed by the monitoring system during peak sunlight hours; it then uses any leftover solar power to charge the battery, ensuring that the system has enough juice to run through the night and on cloudy, rainy days. Since the system nodes are all mill watt-level power consumers, a 30W solar panel can be charged for more than 10 days in a single day, so the use of solar power can meet the system's power supply requirements [26].

The actual power used by the solar panel. The battery adopts a lithium iron phosphate battery with a small size, long life, good environmental compatibility, high-temperature resistance, suitable for fast charging, and high safety factor, the battery capacity calculation formula:

$$B_c = APN_L T_0 / C(Ah) \tag{3.1}$$

In the formula, A is the safety factor; P is the average power consumption, which is the working current multiplied by the daily working hours; N_L is the longest continuous rainy day; T_0 is the temperature correction coefficient; C is the depth of discharge of the battery.

3.3.2. Sensor Node Hardware. Figure 3.2 depicts the sensor node's physical design. The ZigBee module may be expanded with the addition of sensors, cameras, solar power, LED indicators, and buttons because of its CC2430 chip and straightforward peripheral circuit [27]. The CC2430 is an improved ZigBee system-on-a-chip (SOC), incorporating a CC2420RF receiver with high sensitivity and strong anti-interference, as well as a high-performance low-power controller, a direct memory access (DMA) controller, a programmable watchdog timer, an AES-128 security coprocessor, and a 14-bit analog-to-digital converter (ADC). The sensor node is solar-powered, which is a key technical indicator for both small solar panel area and low power consumption.

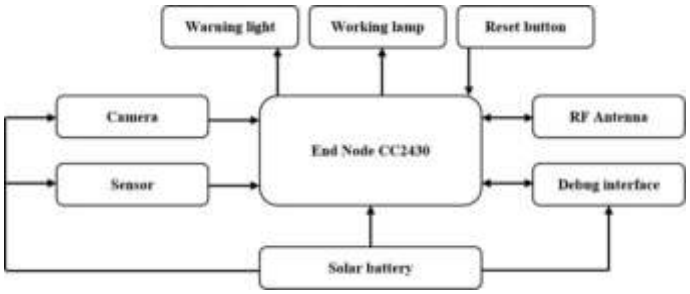


Fig. 3.2: Terminal node structure diagram

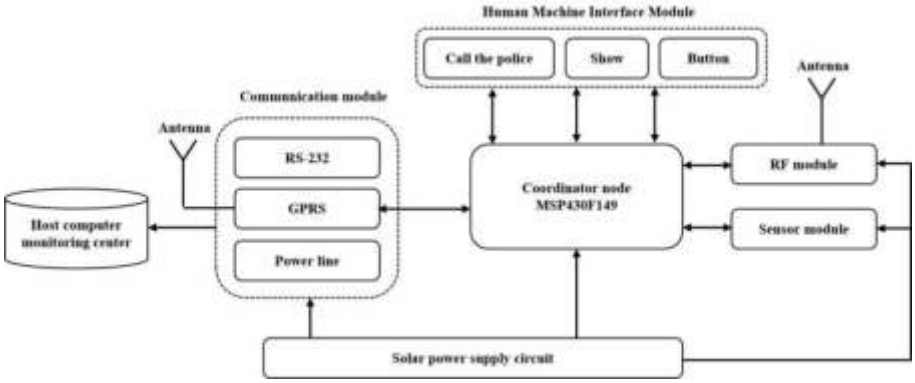


Fig. 3.3: Coordinator node structure diagram

To further reduce power consumption and maximize solar energy harvesting, the node implements a sleep/wake mode, allowing for user-defined data collection intervals and an automatic transition to a low-power state once data transmission is complete [28, 29].

3.3.3. Coordinator node hardware. Figure 3.3 depicts the coordinator node, which employs a 16-bit RISC mixed-signal processor, the MSP430F149 microcontroller. This microcontroller is capable of receiving data, processing it, storing it, displaying it, and transmitting it thanks to its powerful functions, ultra-low power consumption, and abundant peripheral resources. Receive data from other wireless sensor nodes via the radio frequency module, process the data according to a certain algorithm, and send the data to the upper computer data centre via GPRS at the set frequency; when the collected data exceeds the set threshold, it can automatically alarm and send alarm information to relevant personnel via GSM short messages. When the amount of information sent is relatively large, it can also transmit the data via GSM short messages.Finally, when the amount of information sent is relatively large, it can also transmit the data via the machine keeps its original RS-232 serial port for use in file downloads and debugging. If the data collected by a terminal node exceeds the set value or is not in the network, it will alarm and display the relevant information on the LCD screen, and the coordinator node receives the wireless signal sent by the terminal node through the CC2430 module, stores the data in the RAM of the microcontroller, and sends the data to the upper computer regularly.

3.4. System software design.

3.4.1. Software design of lower computer. The software design of the lower computer adopts the cross compiler and debugger of IAR Embedded Workbench7.30B to develop, debug and realize the wireless communication of data, both the coordinator node and the sensor node need to load the TI protocol stack,

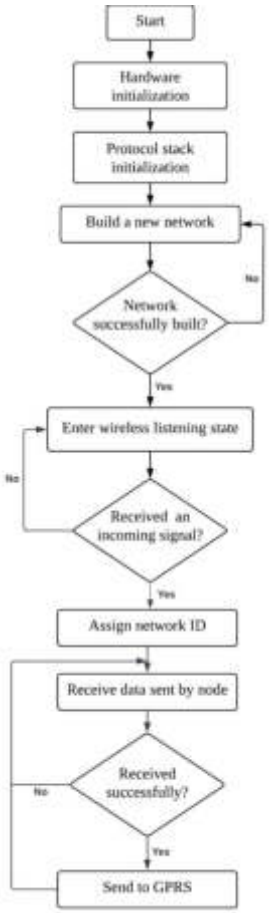


Fig. 3.4: System software flow chart 1

and write application programs at the application layer according to different functions, the coordinator node is responsible for network establishment, automatic networking, GPRS communication, interruption response, etc. From a structural perspective, the terminal node primarily realises the acquisition of sensor data and the wireless transmission of data, and their software architecture is identical, with the primary difference lying in the interaction of the protocol stack, application programme, and operating system. Figure 3.4 depicts the coordinator node's flowchart. First, the coordinator node's MSP430F149 microprocessor, radio frequency module, and liquid crystal are initialized; next, the ZigBee protocol stack is initialized, the interrupt is opened, and the process of establishing a new network begins; finally, the coordinator's physical address, the new network ID number, the channel number, etc. are displayed via the serial port if the network has been successfully established. If the join is successful, the coordinator will begin to receive data from the terminal node. Otherwise, it will attempt to rejoin until it is successful. Afterward, the microcontroller analyses the data that has been wirelessly transmitted to the lab for testing through GPRS or to the appropriate individuals via SMS. The procedure at a sensor node is depicted in Figure 3.5: Like the coordinator, each sensor must first power on and initialize CC2430 before sending an application signal to join the coordinator's network and waiting for a response before receiving an address from the coordinator [30].

If the work light of the terminal node is lit if the join is successful, and the work light of the join fails is off, the node will reapply for joining. After successfully joining the network, the terminal node enters the sleep state, and the terminal node collects sensor data at regular intervals. If the data is normal, the data is packaged and sent to the coordinator by cyclically calling the SendData() function, and the successfully sent

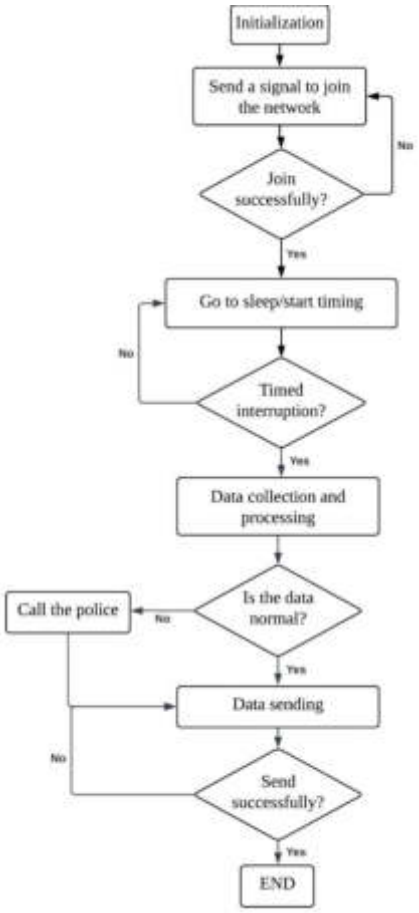


Fig. 3.5: System software flow chart 2

node will re-enter the sleep state; When the data is abnormal, the alarm light will start flashing and the sensor will collect data again and send the data to the coordinator twice. If the transmission is unsuccessful, the sensor immediately re-collects the data again and sends it to the coordinator.

3.4.2. PC software design. The host computer is programmed with Visual Basic 6.0; the prototype system only uses temperature and humidity sensors for data acquisition, and the data received by radio frequency is transmitted to the host computer monitoring center database. Mainly realize the following functions:

1. Real-time data display, the data collected by the terminal sensor is displayed in the form of a real-time curve, and the specific value and time are displayed in the table column on the upper left side, which is refreshed every 3s.
2. Historical record query, the system can query historical records according to different nodes and different periods, and display them in a certain order.
3. System security, when the system is abnormal or faulty, it can automatically save relevant data and can set thresholds for sensors, and automatically alarm when the limit is exceeded.
4. Remote monitoring, the monitoring system can realize remote login through a GPRS connection to Ethernet, and realize remote query and control.

4. Results and Discussion. The prototype undergoes in-depth, objective laboratory testing to evaluate the system's performance metrics. The real temperature and humidity are measured to within 0.1 degrees Celsius (using a conventional thermometer) and within 2% (using a relative hygrometer). Table 4.1 displays

Table 4.1: Test data

RSSI	Measure temperature / ° C	Actual temperature / ° C	Measure humidity %	Actual humidity %
0×85	19.2	18.89	41.23	41.60
0×73	18.8	18.86	42.54	42.98
0×6A	20.1	20.16	43.24	44.62
0×56	19.5	19.42	42.72	43.16
0×48	19.8	19.84	46.36	48.02
0×32	20.4	20.38	45.84	46.24

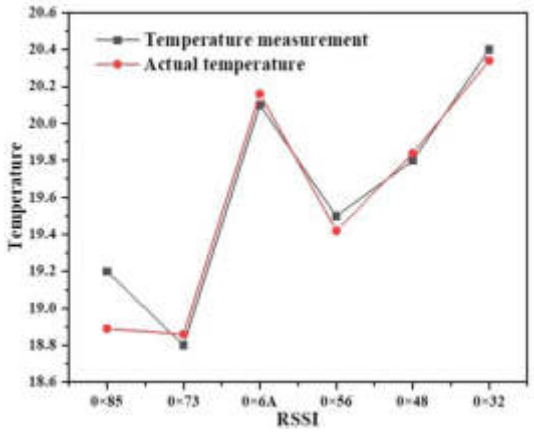


Fig. 4.1: Temperature comparison analysis diagram

the wireless communication distance, the temperature and humidity values collected and sent by the sensor nodes, and the signal strength indicator value received by the coordinator node.

Figure 4.1 shows that the system's greatest temperature difference is 0.31°C and that its maximum humidity difference is 1.38%. Both temperature and humidity thresholds may be configured, with a text message delivered to a mobile device when the measured value rises above the threshold. For instance, the maximum allowable temperature is 30.0 °C and the maximum allowable humidity is 60%. Data is captured every 10s and transferred via GPRS to the monitoring centre; when the threshold is surpassed, a text message reading "Warning! the value is: Temperature: 31.6 °C; Humidity: 41.5%" is delivered.

Two state-of-the-art methodologies were used to evaluate the efficiency of fault location using power IoT in the power system. The superiority of the proposed approach for fault finding is made abundantly clear in Figure 4.2. The suggested method's error increased with increasing fault resistance because it was more affected by fault reactance than the hybrid fault location technique. Even though the location inaccuracy was large when employing the hybrid fault location approach, the issue of high-resistance grounding was ultimately handled. The fully adaptive fault location approach offers fast calculation speeds but poor location performance. However, the method proposed in this study takes into consideration a number of characteristics that improve fault localization. These variables include the node, connecting line, fault type, fault length, and fault resistance. To ensure the safe and dependable operation of power transmission facilities, the operational staff can make choices and issue orders based on the analytical findings of the information management system.

5. Conclusion. In this paper, an Internet of Things-based system is presented for the informatization of transmission line monitoring. The purpose of this effort is to improve upon the weaknesses of the existing power monitoring setup. The system is based around a 16-bit microcontroller that uses very little power and

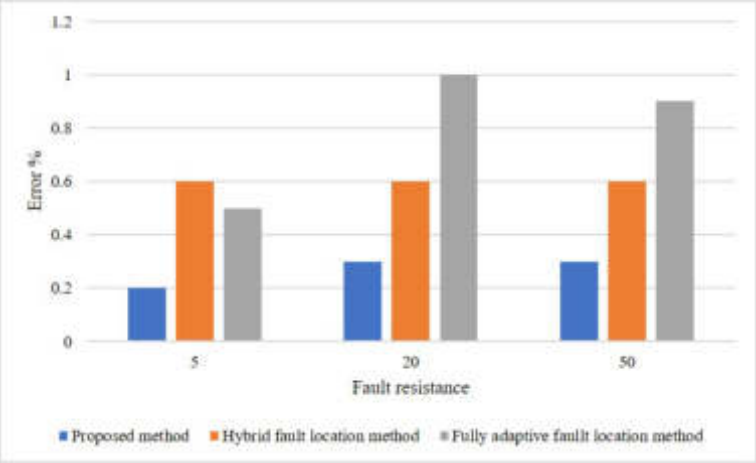


Fig. 4.2: Fault resistance comparison of proposed method with existing methods

can monitor a wireless sensor network. The suggested architecture realizes the collection, transmission, and monitoring of high-voltage transmission line characteristics while also integrating monitoring, display, alarm, communication, and other services into a single package. Accuracy of the method depends on the precision of both the smart control platform system and the measurement data. Future study will investigate how the method proposed in this paper, which is based on IoT data, may be used to distribution network fault finding technology that relies on a smart algorithm to trust the supply generation. Data collected by the prototype shows that the designed system has low measurement error, with most of the error stemming from the sensor itself and almost none coming from the network transmission process. Additionally, the system operates smoothly and reliably, allowing it to fulfill its intended purposes. With this kind of real-time monitoring and early warning of disaster in place, the power system will be better able to withstand or at least mitigate the effects of large natural disasters. Our future work will focus on lowering the price of our power transmission line monitoring technology and increasing its dependability.

6. Acknowledgments. Technology Project of State Grid Electric Power Space Technology Company Limited, Project Name:Research on 3D Delicacy Management Technology of main power grid based on laser scanning, Project No.: 1600 /2020-44003B.

REFERENCES

[1] Ahmad, S., Afzal, M. M. , *Development and effect of fog computing on the big data analysis for iot devices in 5g era. International Journal of Advanced Science and Technology, 29(4), 7033-7040, 2020.*

[2] Kumar, P., Gupta, G. P., Tripathi, R. , *A distributed ensemble design based intrusion detection system using fog computing to protect the internet of things networks. Journal of Ambient Intelligence and Humanized Computing, 12(10), 9555-9572, 2021.*

[3] Kalinin, M. O., Krundyshev, V. M., Sinyapkin, B. G. , *Development of the intrusion detection system for the internet of things based on a sequence alignment algorithm. Automatic Control and Computer Sciences, 54(8), 993-1000, 2020.*

[4] Long, L. , *Research on status information monitoring of power equipment based on Internet of Things. Energy Reports, 8, 281-286, 2022.*

[5] Yu, L., Nazir, B., Wang, Y., *Intelligent power monitoring of building equipment based on Internet of Things technology. Computer Communications, 157, 76-84, 2020.*

[6] Tang, J., Dong, T., Li, L., Shao, L. , *Intelligent monitoring system based on internet of things. Wireless Personal Communications, 102, 1521-1537, 2018.*

[7] Sharma, A. K., Bhandari, R., Aherwar, A., Rimaauskien, U., Pinca-Bretotean, C. , *A study of advancement in application opportunities of aluminum metal matrix composites. Materials Today Proceedings, 26(2), 2419-2424, 2020.*

[8] Huang, R., Yang, X. , *Analysis and research hotspots of ceramic materials in textile application. Journal of Ceramic Processing Research, 23(3), 312-319, 2022.*

[9] Yang, T., Chen, J., Zhang, N. , *Ai-empowered maritime internet of things a parallel-network-driven approach. IEEE Network, 34(5), 54-59, 2020.*
[10] Tang, X. , *Research on smart logistics model based on internet of things technology. IEEE Access, PP(99), 1-1, 2020.*
[11] Bouzidi, M., Dalveren, Y., Cheikh, F. A., Derawi, M. , *Use of the iqrf technology in internet-of-things-based smart cities. IEEE Access, PP(99), 1-1, 2020.*
[12] Yuvaraj, N., Srihari, K., Dhiman, G., Somasundaram, K., Masud, M. , *Nature-inspired-based approach for automated cyberbullying classification on multimedia social networking. Mathematical Problems in Engineering, 2021, 1-12, 2021.*
[13] Khoa, N. M., Cuong, M. V., Cuong, H. Q., Hieu, N. T. T. , *Performance comparison of impedance-based fault location methods for transmission line. International Journal of Electrical and Electronic Engineering Telecommunications, 11(3), 234-241, 2022.*
[14] Iurinic, L. U., Herrera-Orozco, A. R., Ferraz, R. G., Bretas, A. S. , *Distribution systems high-impedance fault location: A parameter estimation approach. IEEE Transactions on Power Delivery, 31(4), 1806-1814, 2015.*
[15] Kong, X., Xu, Y., Jiao, Z., Dong, D., Yuan, X., Li, S. , *Fault location technology for power system based on information about the power internet of things. IEEE Transactions on Industrial Informatics, 16(10), 6682-6692, 2019.*
[16] Feng, X., Qi, L., Pan, J. , *A novel fault location method and algorithm for DC distribution protection. IEEE Transactions on Industry Applications, 53(3), 1834-1840, 2017.*
[17] Zhang, J. , *Distributed network security framework of energy internet based on internet of things. Sustainable Energy Technologies and Assessments, 44, 101051, 2021.*
[18] Lei, T., Lv, F., Liu, J., Feng, J. , *Research on electrical equipment monitoring and early warning system based on Internet of things technology. Mathematical Problems in Engineering, 2022, 1-12, 2022.*
[19] Zhao, J., Yue, X. , *Condition monitoring of power transmission and transformation equipment based on industrial internet of things technology. Computer Communications, 157, 204-212, 2020.*
[20] Kong, P. Y., Song, Y. , *Joint consideration of communication network and power grid topology for communications in community smart grid. IEEE Transactions on Industrial Informatics, 16(5), 2895-2905, 2020.*
[21] Whitehouse, K., Blaylock, R., Makleff, S., Lohr, P., *"it's a small bit of advice, but actually on the day, made such a difference. " perceptions of quality in abortion care in england and wales. Reproductive Health, 18(1), 1-9, 2021.*
[22] Hajahamed, M. A., Hawa, M., Shamlawi, L., Alnaser, S., Abualnadi, D. , *Cognitive radiobased backup protection scheme for smart grid applications. IEEE Access, PP(99), 1-1, 2020.*
[23] Guo, E., Jagota, V., Makhatha, M. Kumar, P. , *Study on fault identification of mechanical dynamic nonlinear transmission system. Nonlinear Engineering, 10(1), 518-525, 2021.*
[24] Vorona, A. , *Global factors influencing the innovative development of the national economy. Economics Finances Law(3(1)), 32-36, 2020.*
[25] Ortiz, G. R., Cespedes-Panduro, B., Shahrtash, S. A., Rahimi, F., Sandi, S., Arias-Gonzáles, J. L., ... Akhavan-Sigari, R. , *DFT assessments of BN, AlN, and GaN decorated carbon cage scaffolds for sensing the thiamazole drug. Diamond and Related Materials, 135, 109800, 2023.*
[26] Song, S., Dong, C., Qin, K., Hou, X., Zhao, X., Tian, X., Wang, H. , *Research on Time Synchronization and Spatial Calibration Algorithms of Radar and Camera. In Journal of Physics: Conference Series, 2428(1) 012044, 2023.*
[27] Kou, R., Lian, S. W., Xie, N., Lu, B. E., Liu, X. M. , *Image-based tool condition monitoring based on convolution neural network in turning process. The International Journal of Advanced Manufacturing Technology, 119(5), 3279-3291, 2022.*
[28] Sitharthan, R., Vimal, S., Verma, A., Karthikeyan, M., Dhanabalan, S. S., Prabaharan, N., ... Eswaran, T. , *Smart microgrid with the internet of things for adequate energy management and analysis. Computers and Electrical Engineering, 106, 108556, 2023*
[29] Mohammed, K., Abdelhafid, M., Kamal, K., Ismail, N., Ilias, A. , *Intelligent driver monitoring system: An Internet of Things-based system for tracking and identifying the driving behavior. Computer Standards Interfaces, 84, 103704. 2023.*
[30] Yang, B., Tang, J., Dong, X., Li, S., Gu, R., Hao, J. , *Power Inspection Design by Internet of Things and RFID Technology in Smart City. Microprocessors and Microsystems, 96, 104510, 2023.*

Edited by: Pradeep Kumar Singh
Received: Mar 15, 2023
Accepted: Jul 28, 2023


Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2023 SCPE. Volume 24, Issues 2, pp. 127–137, DOI 10.12694/scpe.v24i2.2146

COMPUTER HARDWARE AND NETWORK DATA TRANSMISSION BASED ON INTERNET OF THINGS COMMUNICATION TECHNOLOGY

LING WANG *

Abstract. In order to meet the requirements of computer hardware and network data transmission security, a research based on Internet of Things communication technology is proposed. The main content of this research is the research based on the communication technology of the Internet of Things, through the description of the communication protocol of the Internet of Things, the system hardware design and implementation methods are used, and finally the research method based on the communication technology of the Internet of Things is constructed through experiments and analysis. The core technologies of 5G connectivity are being used to construct the IOT. As a result, the IOT might gain momentum. The experimental findings demonstrate that the delays are all within 200 Ms. When the message size is short (within 1KB), the transmission of diverse hardware is average, and the transmission quality standards of QoS1 are fulfilled. The transmission quality standards of QoS1 can match the communication reliability and security needs of the Internet of Things. This article evaluates the performance of data transfers with lengths of 20 byte, 30 byte, 50 byte, and 70 byte, respectively. This paper evaluates the efficiency of Wi-Fi access configuration by sending data packets of varying sizes i.e., 10 bytes, 30 bytes, 50 bytes, and 70 bytes over a distribution network. The results show that, on average, the network takes 0.6692s, 1.3546s, 2.8600s, and 4.7319s to deliver each packet, with success rates of 100% in each case. The system's increased network distribution efficiency is observed from the experimentation. The research based on the Internet of Things communication technology can meet the needs of computer hardware and network data transmission security.

Key words: Internet of things; Communication technology; Computer hardware; Network data transmission.

1. Introduction. With the continuous advancement of science and technology, it has greatly promoted the emergence and development of the Internet. The role of computer hardware and network data transmission is becoming more and more important. Increasing the use of computer equipment and communication technology on the Internet can improve the remote operation level of telecommunication network technology, fully evaluate the use of computer equipment and communication technology, and meet the requirements of reducing costs, meet the needs of self-improvement, the efficiency of data transmission and the level of network transmission are steadily improved, and the effective use and reconfiguration of network resources are realized [1].

The application of computer hardware in current applications is mainly reflected in the application of basic computer equipment such as cameras, RFID scanners, etc., which is done in basic hardware applications [2]. Computer terminal detectors help complete the collection and transmission of basic data. Through its application, all things can be effectively connected to the Internet, the sensor is mainly composed of power management module, sensor module and LED display. In addition, access control sensors also account for a large proportion of applications. Gateway sensors act as routers to transmit signals from network endpoints. The gateway is arguably the sensor that acts as a bridge, ensuring that the communication network can be effectively connected and ultimately feel in control.

With the continuous deepening of the application of network technology, in order to effectively improve the quality of computer network services, people's requirements for network service quality are getting higher and higher [3]. Achieve service quality improvement and standardized improvement processing, it is necessary to continuously improve various implementation conditions and implementation time, effectively improve the quality of network services. If the network is not secure, it will seriously affect the operation of the entire network, and the security of network services is an important prerequisite for network operation. In order to prevent the outside world from causing serious interference to the operation of the entire network, it is necessary to ensure the reliability of the server. For example, the servers of some enterprises are located abroad, whether

*College of Computer Science, Baicheng Normal University, Baicheng, Jilin, 137000, China (lingwang215@gmail.com)

people's private data has been protected, their security must be studied. If an operation failure occurs during the operation, the user can take timely measures to protect the information, and the monitoring system can obtain the information in time and return it to the user in real time, reducing the degree of loss. One of the IOT's initial technologies is communication technology. Dissimilar categories of communication systems can be used to connect the IOT, and the most recent advancement in 5G communication has distinct compensations.

1.1. Problem Statement. The use of Internet of Things (IoT) communication technology in computer hardware and network data transmission has introduced several research problems. One of the major issues is the reliability of network transmission in complex IoT systems. The large volume of data generated by IoT devices can cause network congestion and latency, leading to transmission errors and delays. Additionally, the use of different communication protocols and standards in IoT devices can create interoperability issues and hinder data transmission. Another problem is the security of data transmission in IoT systems. Due to the large number of connected devices and the lack of standard security protocols, IoT systems are vulnerable to cyber attacks and data breaches. Insufficient data protection, insecure communications, and data storage are some of the significant challenges for IoT privacy and security. To address these research problems, several approaches have been proposed. One approach is to use edge computing to reduce network congestion and latency by processing data closer to the source. Another approach is to develop standard security protocols and frameworks for IoT devices to ensure data confidentiality and integrity. The use of IoT communication technology in computer hardware and network data transmission has introduced several research problems related to network reliability and security. Addressing these problems requires the development of new approaches and standard protocols to ensure the secure and efficient transmission of data in IoT systems.

1.2. Motivation and Contribution. For the Internet of Things to be used on a broad scale, more study has to be done to understand its architecture, communication mechanisms, and networking methods, as well as the information interaction of the complex interactions between devices [4]. When it comes to the Internet of Things (IoT), existing communication architectures typically rely on the cloud platform, which serves as middleware and connects IoT devices together. While this solution promotes the diversity of heterogeneous devices from application developers or device manufacturers, it also increases the platform's complexity [5, 6]. In the 5G era, the national growth strategy clearly guides the development of IOT technology and applications. IoT technology should be aggressively promoted throughout a wide range of businesses, development fields, and application domains. Our goal is to increase communication, traffic, and resource pooling. We also encourage new approaches, new ideas, and new ways of integrating data. As 5G wireless networks become more widely available, the Internet of Things will evolve. As a result, we take a look at the scientific strategies that have been implemented to modernize and improve age-old industries. Meanwhile, IoT applications can boost the marketability of useful data and services.

1.3. Organization of the article. The remaining article is structures as: Literature survey of this article is presented in section 2 of the article followed by Research on Internet of Things Communication Technology briefed in section 3. Section 4 presents the analysis of results and the final conclusion is detailed in section 5.

2. Literature Review. In the current social development, the rapid development and progress of the level of science and technology, people are more and more inseparable from the use of computers in their daily lives, because computers can be used for communication to process some necessary information, moreover, daily communication can be carried out through the computer, which makes the weight of computer communication improved to a certain extent [7]. At present, precisely because of the development and progress of the communication industry, in terms of network, the situation of network saturation has gradually appeared, therefore, more and more advanced communication technology is needed to solve the current pressure of network communication, so that all walks of life can achieve smooth communication and interaction with the help of communication technology, in addition, communication interference in the service can be avoided, and the pressure on the computer in terms of communication can also be correspondingly weakened and controlled. In addition, the technology can also enable the computer to establish a more complete platform for network monitoring to realize real-time monitoring of the background work process. Under the influence of the continuous development of science and technology, the use of communication technology on the Internet is limited and the network is relatively wide. Internet communication and communication technologies can be established.

The Internet of Things is the age of applications. Today, there is a lot of attention on the Internet, and most executives use the opportunity to actively explore and discuss the Internet. The interconnection of communication technologies and mobile payments has contributed to the development of various communication industries. In short, the potential for identifying and using communication networks is relatively large, and thus can be actively promoted on the Internet [8]. With the widespread use of artificial intelligence technology, modern information technology, with the advancement of technology, the Internet of Things has been gradually introduced, so that the actual operation of the objects used in the Internet of Things can be remotely controlled through communication and information technology, use remote rice cookers and hot water at home with the help of communication technology. When you're at work, use the app on your phone to control your home remotely. When you get home from get off work, the hot water will boil in the water heater, which not only saves time, but also improves people's living standards. In view of the above problems, in order to meet the needs of computer hardware and network data transmission security, the author proposes a research based on Internet of Things communication technology [9]. The realization of the Internet of Things technology takes the Internet as the carrier, and realizes the mutual transmission of data and information through the Internet, which is no longer limited to the limited physical space, in fact, the nature of the information transmitted is mainly virtual information and analog signals, however, IoT technology can use signal transmission to control real things, build local area networks, and connect various types of computer communication networks, ensure that data transmission and signal transmission are faster and faster, expand the controllable range of IoT technology, give full play to the advantages of computer communication networks, and create a remote management and control network that integrates informatization and intelligence.

As the Internet of Things scope and number of device accesses grow, so does the volume of data processing, which raises the threshold for system dependability, stability, and concurrent data processing [10]. It is imperative that research be conducted on the Internet of Things's design, communication, and networking techniques, as well as the information interaction of the complex relationships between devices [11], if it is to be widely adopted. Existing IoT communication architectures are typically based on cloud platforms, which act as middleware and connect IoT devices. This solution encourages application developers and device manufacturers to create heterogeneous devices due to the complexity of the platform, but a standard protocol is required for the communication of these devices. The MQTT protocol is the most prevalent among the numerous IoT communication protocols [12]. The MQTT protocol's topic subscription / message delivery method allows the monitoring centre and smart gateway to remotely manage and control papermaking equipment located in several locations through a cloud-based MQTT server. A huge amount of historical data saved on a cloud server may be mined and processed using data mining and information technology to aid in the making of scientific judgments, the improvement of production efficiency, and the reduction of expenses. The system's reliability and stability were shown by high-frequency vibrator testing [13]. Based on the MQTT protocol and the theoretical basis of Wi-Fi technology, this article provides a strategy for optimizing Wi-Fi distribution networks, a mechanism for communicating between terminals, and a platform for testing the system. As the results of the experiments show, the MQTT-based communication method for IoT terminal equipment suggested in this work can reliably and flexibly execute the equipment's communication function and fulfill the communication needs of the IoT system.

3. Research Methods.

3.1. Research on Internet of Things Communication Technology.

3.1.1. The main architecture of the Internet of Things. The realization of the function of the Internet of Things needs to rely on a perfect logical structure, which generally includes an information perception and control module, an information transmission module and an information application module [14].

Information application module. This module, also known as the IoT application layer, covers a variety of application software, because the Internet of Things system has huge data resources, combined with big data technology, cloud computer technology, etc., it can control and manage actual objects on the network platform.

Information transfer module. This module is also known as the transport layer. At this stage, the network forms used for Internet of Things information transmission include mobile communication networks, local area networks, and the Internet, data transmission between different networks relies on communication technology,

connect different types of information management and sensing nodes, build an IoT information transmission and resource management and control platform, inorderto ensure the stability and security of the information transmission of each module [15].

Information perception and control module. This module, also known as the perception layer, is composed of different types of controllers and sensors, its main function is to perceive the target object information within the scope of the Internet of Things, at the same time, the acquired data information is converted into a common communication format, and then data transmission is completed based on wired or wireless networks [16].

The perception layer is one of the important components to ensure the normal operation of the IoT system, its function is to perceive and collect data information in the area, and use the IoT control box to complete the data information transmission between different gateways, establish a stable connection relationship with multiple local area networks, exchange information with each sensor, execute professional program algorithms in the system to analyze data information, and complete the processing of related data. The data information acquired by the perception layer is the basis for the operation of the computer network, and many tasks need to be based on this, because the acquired data information has good accuracy and comprehensiveness, and the network control box completes data transmission, protocols and groups. network transformation, etc., further expand the information surface that the computer network can cover [17]. The system can select corresponding analysis methods according to different data types, which makes data processing work smoother and more efficient, and effectively improves the processing speed of system data information, in the actual use of the Internet of Things, a reasonable method is adopted to allocate limited Internet resources, relying on the current communication network, with the continuous improvement and development of the Internet of Things technology, more and more functions and scope of application can be realized, especially in the field of logistics, the IoT perception layer can identify huge logistics data information, and it is transmitted to the computer communication network system, which can complete the processing of huge data information in a relatively short period of time, effectively improve the work efficiency in the field of logistics, and promote the further development of the logistics industry.

The Internet of Things transport layer relies on the existing communication network to strengthen the connection with the information perception network, communication network and broadband through computer communication methods in the form of optical fiber width and wireless network. With the continuous improvement of the level of modern science and technology, the Internet of Things technology has also achieved greater breakthroughs. On the one hand, it promotes the further improvement of existing related technologies, on the other hand, some communication nodes with strong independence and extensibility have also been produced, and these new communication nodes are fundamentally different from general computer network communication nodes, which brings huge challenges to the operation of computer network communication systems [18]. In addition, the computer communication network mainly uses the Internet of Things to realize the management and control of each node, which can significantly improve the speed of the Internet of Things. At present, during the joint operation of multiple computer networks, there will be problems of mutual influence. In order to speed up data processing, it is necessary to create a sound operation platform and fully understand the needs of different industries, actively carry out the development of computer communication network monitoring functions, build a reasonable industrial development system, set up cloud integration connection ports, and complete the disclosure and sharing of information in various fields on the operating platform, create a complete Internet application platform based on cloud computing technology, and further strengthen the operation effect of computer communication network.

3.1.2. IoT Communication Protocol. The MQTT protocol is one of the earlier communication protocols of the Internet of Things, it is a lightweight transmission protocol based on a subscription and publishing mechanism to solve the communication between devices in poor network conditions [19]. Under a topic that publishes a message to the server through an MQTT client, any MQTT client that subscribes to the topic can receive the message, and the client can receive messages under multiple topics by subscribing to multiple topics. MQTT is a TCP-based communication protocol and supports TLS/SSL encryption, and its minimum packet size is only 2 bits. Due to the complexity of the industrial site environment, the instability of the network and the need to access a large number of local sensors, a high-quality, dependable, and low-bandwidth data exchange protocol is required for the industrial application scenario of cloud-side communication. Due to its

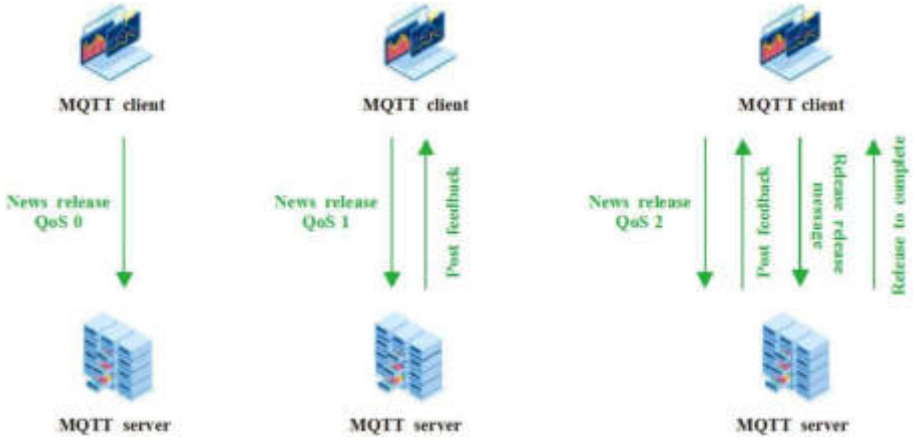


Fig. 3.1: MQTT protocol working mechanism

small message size, publishing and subscription-based decoupled communication modes, extensive QoS options, and LastWill mechanism, the MQTT protocol is highly suited for application scenarios with limited hardware resources and bandwidth. Therefore, the MQTT protocol has become a viable option for the industrial IoT communication system.

MQTT provides a will mechanism and three flexible mechanisms to meet varying data transmission quality requirements. In the Internet of Things, QoS 0 or QoS 1 can be used for uploading and analyzing sensor data to the cloud. QoS 2 can be used to guarantee the accuracy and timeliness of transmitted data for vital parameters and sensors. As depicted in Figure 3.1, when the sensor or gateway is abnormally offline, other clients can also be notified in a timely manner via the will mechanism. Commonly used IoT information and communication technology (ICT) protocols, compared to other IoT ICT protocols, the MQTT protocol provides more quality-of-service choices in the service mode that supports publish and subscribe, and because of its smaller message size, it has lower requirements for bandwidth [20]. At the same time, the TCP-based message transmission of the MQTT protocol ensures the reliability of the message transmission process. Therefore, the author uses the MQTT protocol as the on-site cloud-side communication protocol to build the system.

3.2. System Design.

3.2.1. System Hardware Design and Implementation. According to the architecture of the IoT system, the author designs and implements the IoT communication system of MQTT. The system includes perception layer, transport layer, platform layer and service layer. The perception layer includes on-site vibration sensors and video image collectors, data collectors and intelligent gateways; Sensors and data collectors in the perception layer form edge nodes, and edge gateways form the edge of the system, which communicates data with the cloud through the MQTT protocol through the transport layer, among them, the transmission layer adopts 4G network to realize the data communication between the field and the cloud; The platform layer adopts Alibaba Cloud server, deploys server application software, realizes data storage, analysis, processing and forwarding on the edge site, and analyzes user subscription services; The service layer is based on technologies such as Qt and Paho, and develops application software on the PC client to realize user theme service subscription and data visualization [21].

The ZM-YB40-V/A sensor is a three-wire instrument that can produce either an analogue voltage or current signal, making it a prototypical vibration sensor [22]. In order to gather sensor data, the author used the low-power, multi-functional acquisition board's central CPU and the analog-to-digital conversion module. Figure 3.2 depicts the three components that make up this system's software architecture: an edge data collection module, a cloud-edge collaborative communication module, and a user subscription service module.

The edge data acquisition subsystem includes on-site data acquisition and data processing and analysis modules, the data acquisition of the vibration sensor is realized by using the STM32 on-chip AD function [23].

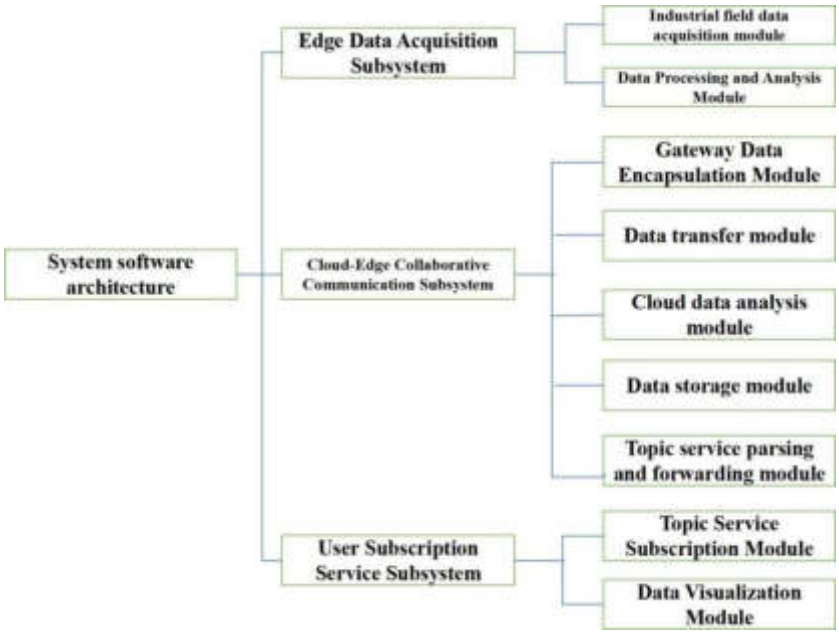


Fig. 3.2: Functional system software architecture

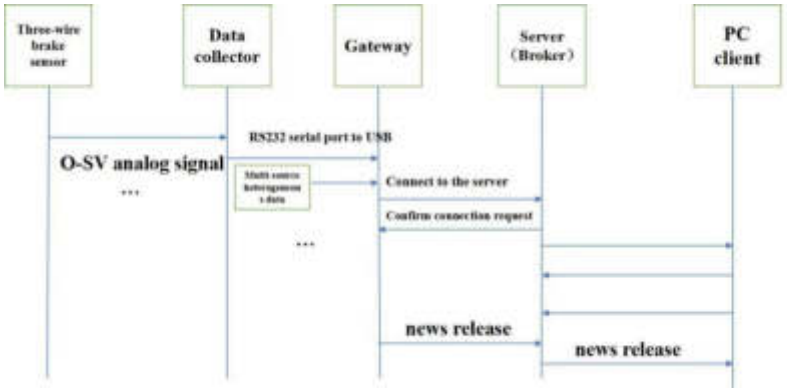


Fig. 3.3: Information flow

The data collector is connected with the gateway through RS232, and transmits data by serial communication. Figure 3.3 depicts the informational flow throughout the system. Images and data from vibration sensors are taken at the test location, but only the images are sent. The system's features include cloud data analysis, subject subscription services, and the collecting, processing, and transmission of multi-source heterogeneous data from industrial locations.

3.2.2. System Testing and Analysis. In the delay test, the time for data transmission from the gateway to the cloud is t_1, and the time for the cloud server to parse, store, and queue the message is Δt, the time of transmission from the cloud to the subscription client is t_2, and the total time t_3 for the message to be transmitted from the gateway to the subscription terminal can be obtained from formula (3.1):

$$t_3 = t_1 + \Delta t + t_2 \tag{3.1}$$

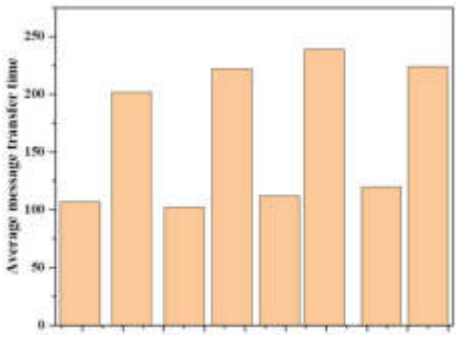


Fig. 3.4: Average delay in sending data packets on PC

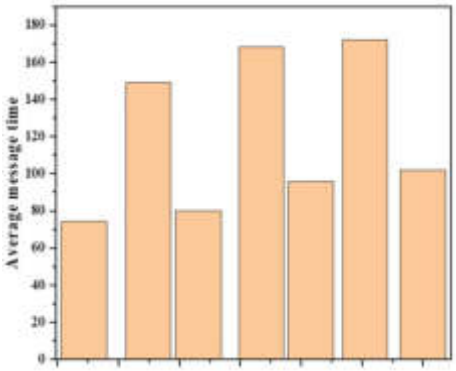


Fig. 3.5: Average delay in sending data packets on Raspberry Pi

Taking into account the actual situation, the time data obtained by the author's test are the total time (i.e. t3) of data transmission from the gateway to the subscriber, and the standard deviation of the data can be obtained by formula (3.2):

$$std.Deviation = \sqrt{1/n \sum_{i=1}^{n} (t_i - t_\mu)^2} \tag{3.2}$$

Among them, n is the total number of tests in each group, t_i is the single measurement delay of each group, and $t_\mu$is the average delay of the group. In order to determine the latency performance of the server when the client node and message load volume gradually increase, a latency test needs to be performed, the test tool is an open-source script, the script can simulate a large number of clients connecting to the server concurrently, and can set parameters such as QoS and packet volume [24,25]. Figure 3.4 depicts the mean time taken to transmit 10 data packets with varying client counts and quality settings. Figure 3.5 displays the median time taken for transmitting 10 data packets with varying client counts and quality settings on a RaspberryPi.

4. Analysis of results. Table 4.1 shows the average delay, minimum delay and maximum delay of 10 packets tested by a single client under different quality requirements and different packet volumes under different hardware terminals [22, 23].

It can be seen from the results that when the network conditions are good, the message is successfully transmitted, and the proportion of successful transmission is also 100%, that is, there is no information loss phenomenon. When the message size is small (within 1 KB), under the quality requirements of QoS1 and QoS2, the average transmission delay of different hardware is within 200 ms. When the message size exceeds

Table 4.1: Message volume test data

Message volume	Hardware	QoS	Average transmission (delay ms)	Standard deviation (ms)	Minimum transmission (delay ms)	Maximum transmission (delay ms)
10Byte	PC	QoS1	100.18	52.84	59.35	216.43
10Byte	PC	QoS1	164.38	49.82	117.91	260.81
10Byte	Raspberry Pi	QoS1	86.49	25.59	62.19	135.15
10Byte	Raspberry Pi	QoS2	153.70	31.85	123.59	219.31
1KB	PC	QoS1	95.424	57.72	61.91	302.49
1KB	PC	QoS2	162.83	42.49	124.861	250.97
1KB	Raspberry Pi	QoS1	89.23	29.19	59.98	159.01
1KB	Raspberry Pi	QoS2	169.59	49.58	127.75	281.54
500KB	PC	QoS1	246.55	156.16	158.99	652.61
500KB	PC	QoS2	376.20	115.69	189.37	759.82
500KB	Raspberry Pi	QoS1	1602.88	1222.36	398.61	3211.97
500KB	Raspberry Pi	QoS2	2679.88	1797.99	419.97	5983.95
1MB	PC	QoS1	487.96	203.48	326.9	957.18
1MB	PC	QoS2	733.58	224.59	377.67	1136.05
1MB	Raspberry Pi	QoS1	4078.30	1869.05	792.10	6883.91
1MB	Raspberry Pi	QoS2	5630.02	3167.08	831.82	11533.821

Table 4.2: Test outcomes of proposed method

Performance parameter	20 Bytes Message length	30 Bytes Message length	50 Bytes Message length	70 Bytes Message length
Average time	0.67	1.4	2.9	4.8
Shortest time	0.48	1.3	2.2	2.2
Longest time	0.83	2.4	3.5	8.9
Success rate	98.56%	97.08%	99.76%	99.54%

1 KB, the communication delay performance of Raspberry Pi deteriorates rapidly, and it is also significantly better than QoS2 in the case of QoS1. The change of the communication delay under the PC side is small, which has a great relationship with its hardware resources. When the message size increases to 500 KB and above, the transmission time increases significantly, and the transmission delay under QoS2 quality on the PC reaches 600-800 ms. For Raspberry Pi with poor hardware resources, it takes an average of 5.6s to transmit 1 MB of data under QoS2, and according to Table 4.1, in the whole test process, the standard deviation of communication delay with larger message volume is larger, indicating that the transmission speed is unstable. The worst case in the test, where the Raspbery Pi transfers 1MB of data with QoS 2, took 11.5s.

The results of the proposed configuration's performance tests are shown in Table 4.2 and Figure 4.1. In order to evaluate the efficacy of the setup technique proposed in this paper's terminal apparatus, the supply time of the transmitted information is 20, 40, 50, and 70 bytes. In all four test groups, the configuration success rate for the IoT terminal equipment devised in this study approaches 100 percent, demonstrating the method's high success rate. The average time is only 0.78 seconds for 20 bytes and 4.9 seconds for 70 bytes, resulting in a significant decrease in distribution time, an increase in distribution efficiency, and the ability to meet the requirements of rapid and stable distribution. The performance of the proposed system is also evaluated with increasing bytes, namely 25, 40, 60, and 70 bytes, as shown in Figure 4.2. Observations indicate that as bytes increase, the success rate increases to 100 percent. In conclusion, under identical hardware conditions, the number of consumers has little influence on the communication delay of the system. However, under the same hardware resources, increasing the packet size will substantially increase the system's communication latency (especially when the packet size is greater than 1 KB), so this system is suited for the transmission of small

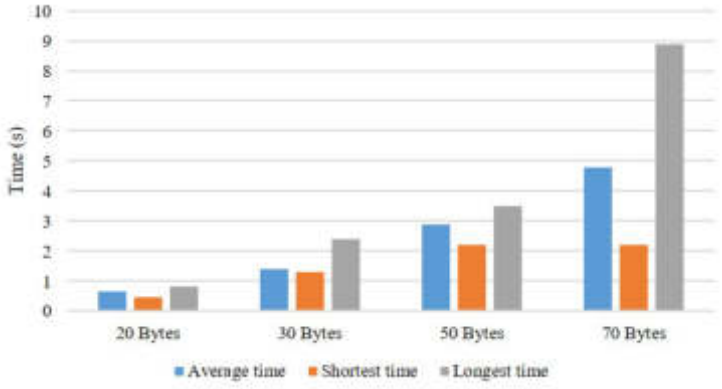


Fig. 4.1: Performance test outcomes of proposed method

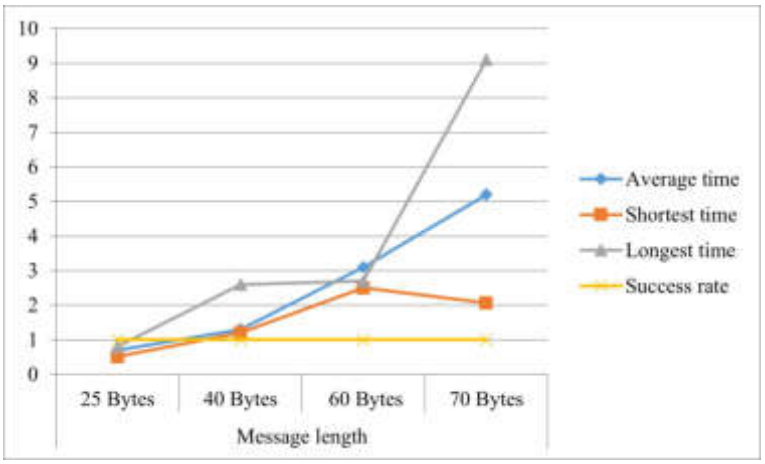


Fig. 4.2: Performance evaluation of proposed method for 25 bytes, 40 bytes, 60 bytes and 70 bytes

data packets and scenarios with high concurrent access. Similarly, the QoS1 transmission quality requirements can satisfy the IoT communication reliability requirements.

5. Discussion. This study investigates the means through which IoT terminal devices can connect to cloud platforms. The suggested method involves the device's domain name first connecting to the cloud platform for identity authentication, and then using that information to connect to the MQTT proxy server. Research the HTTP protocol's data format thoroughly, build an HTTP client within the terminal, and provide a request interface so that the device may connect to the cloud. In this post, we take a look at how Wi-Fi networks are set up and how they code their data. The UDP multicast addresses and data length are developed to facilitate easy, quick, efficient, and inexpensive network connectivity for the device. We have finished the specific implementation on the setup tool and terminal equipment, and have developed a coding network distribution mechanism and a coding protocol. This article establishes a system test platform and finishes testing the terminal device's network access setup technique for functionality and performance. Based on the findings, it is clear that the terminal equipment's suggested way of accessing the network and exchanging data is both practical and effective in meeting the Internet of Things' communication requirements.

6. Conclusion. In order to meet the requirements of computer hardware and network data transmission security, the author proposes a research based on Internet of Things communication technology. The application

of Internet of Things technology can realize the information exchange between people and objects, objects and objects, and then realize the management and control of related objects. The Internet of Things mainly includes three parts: information application module, information transmission module, information perception and control module. With the continuous development of the Internet of Things technology, the perception layer, the transmission layer and the application layer have an increasing impact on the computer communication network, continuous technological innovation must be carried out on the basis of full analysis of market demand, only then can we ensure the perfect integration of IoT and computer communication network. In this article, we take a closer look at the suggested method's underlying coding mechanism and network distribution strategy. Network communication is simplified, accelerated, maximized, and reduced in cost when a multicast address and data length are generated for the device. Based on the results, it is clear that the proposed network access strategy and communication strategy for the terminal equipment is feasible and can meet the communication needs of the Internet of Things system.

REFERENCES

[1] Yang, L., *Research on the utilization of computer network information resources and the reform of college physical education. Journal of Physics: Conference Series, 1915(3), 032060, 2021*

[2] Zhang, N., Xie, H., Li, J., Chen, H., *Application of radar signal processing and image display algorithm based on computer hardware system in intelligence processing. Microprocessors and Microsystems, 81(2), 103747, 2020*

[3] Lu, H., Li, M., Zhang, Y., *Research on optimization method of computer network service quality based on feature matching algorithm. Journal of Physics: Conference Series, 1982(1), 012005, 2021*

[4] Gupta, B. B., Quamara, M., *An overview of Internet of Things (IoT): Architectural aspects, challenges, and protocols. Concurrency and Computation: Practice and Experience, 32(21), e4946, 2020*

[5] Javadpour, A., Wang, G., Rezaei, S., *Resource management in a peer to peer cloud network for IoT. Wireless Personal Communications, 115, 2471-2488, 2020*

[6] Chen, Y., Liu, W., Niu, Z., Feng, Z., Hu, Q., Jiang, T., *Pervasive intelligent endogenous 6G wireless systems: Prospects, theories and key technologies. Digital communications and networks, 6(3), 312-320, 2020*

[7] Fu, N., Zhang, D., *Research on application technology of computer communication and network development based on data mining technology. Journal of Physics: Conference Series, 1982(1), 012132, 2021*

[8] Chien, H., Hori, K., Saito, O., *Urban commons in the techno-economic paradigm shift: an information and communication technology-enabled climate-resilient solutions review:. Environment and Planning B: Urban Analytics and City Science, 49(5), 1389-1405, 2022*

[9] Tan, J., Xia, D., Dong, S., Zhu, H., Xu, B., *Research on pre-training method and generalization ability of big data recognition model of the internet of things. ACM Transactions on Asian and Low-Resource Language Information Processing, 20(5), 1-15, 2021*

[10] Hou, R., Ren, G., Zhou, C., Yue, H., Liu, H., Liu, J., *Analysis and research on network security and privacy security in ubiquitous electricity Internet of Things. Computer communications, 158, 64-72, 2020*

[11] Aman, A. H. M., Shaari, N., Ibrahim, R., *Internet of things energy system: Smart applications, technology advancement, and open issues. International Journal of Energy Research, 45(6), 8389-8419, 2021*

[12] Zhang, W., Ma, F., Ren, M., Yang, F., *Application with Internet of things technology in the municipal industrial wastewater treatment based on membrane bioreactor process. Applied Water Science, 11(3), 1-12, 2021*

[13] Liu, X., Zhang, T., Hu, N., Zhang, P., Zhang, Y , *The method of Internet of Things access and network communication based on MQTT. Computer Communications, 153, 169-176, 2020*

[14] Martin, E. T., Blatt, P., Nguyen, E., Lahr, R., Selvam, S., Yoon, H., *A translation control module coordinates germline stem cell differentiation with ribosome biogenesis during drosophila oogenesis. Developmental Cell, 57(7), 883-900.e10, 2022*

[15] Hsu, W., Kastner, D. B., Baccus, S. A., Sharpee, T. O., *How inhibitory neurons increase information transmission under threshold modulation. Cell Reports, 35(8), 109158, 2021*

[16] Vien, Q. T., Agyeman, M. O., Tatipamula, M., Nguyen, H. X., *On the cooperative relaying strategies for multi-core wireless network-on-chip. IEEE Access, PP(99), 1-1, 2021*

[17] Guo, B., Zang, W., Zhang, R., *Soil salizanation information in the yellow river delta based on feature surface models using landsat 8 oli data. IEEE Access, PP(99), 1-1, 2020*

[18] Azari-Nasab, T., Ghobadi, C., Azarm, B., Majidzadeh, M., *Triple-band operation achievement via multi-input multi-output antenna for wireless communication system applications. International Journal of Microwave and Wireless Technologies, 12(3), 259-266, 2020*

[19] Liu, W., Wu, H., Meng, T., Wang, R., Xu, C. Z., *Aucswap: a vickrey auction modeled decentralized cross-blockchain asset transfer protocol. Journal of Systems Architecture, 117(4), 102102, 2021*

[20] Littman, N., O'Keefe, S. G., Galehdar, A., Espinosa, H. G., Thiel, D. V., *Bandwidth control of loop type frequency selective surfaces using dual elements in various arrangements. Flexible and Printed Electronics, 6(4), 045009, 2021*

[21] Ning, B., Chen, Z., Tian, Z., Wang, X., Li, S., *Joint power allocation and passive beamforming design for irs-assisted physical-layer service integration. IEEE Transactions on Wireless Communications, PP(99), 1-1, 2021*

[22] Ricci, D., Tutton, S., I. M., Ying, M., Blumenthal, D., Eletto, D., *An interdomain helix in ire1 mediates the conformational change required for the sensor's activation. Journal of Biological Chemistry, 296(3), 100781, 2021*
[23] Bedhief, I., Kassar, M., Aguili, T., *Empowering SDN-Docker Based Architecture for Internet of Things Heterogeneity. Journal of Network and Systems Management, 31(1), 14, 2023*
[24] Zhang, M., Ghodrati, N., Poshdar, M., Seet, B. C., Yongchareon, S., *A construction accident prevention system based on the Internet of Things (IoT). Safety Science, 159, 106012, 2023*
[25] Mohammed, K., Abdelhafid, M., Kamal, K., Ismail, N., Ilias, A., *Intelligent driver monitoring system: An Internet of Things-based system for tracking and identifying the driving behavior. Computer Standards Interfaces, 84, 103704, 2023*
[26] Salehi, S., Farbeh, H., Rokhsari, A., *An adaptive data coding scheme for energy consumption reduction in SDN-based Internet of Things. Computer Networks, 221, 109528, 2023*
[27] Ryalat, M., ElMoaqet, H., AlFaouri, M., *Design of a smart factory based on cyber-physical systems and internet of things towards industry 4.0. Applied Sciences, 13(4), 2156, 2023*

Edited by: Pradeep Kumar Singh
Received: Mar 15, 2023
Accepted: Jul 28, 2023

Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2023 SCPE. Volume 24, Issues 2, pp. 139–149, DOI 10.12694/scpe.v24i2.2150

APPLICATION OF ASSOCIATION RULE MINING ALGORITHM BASED ON 5G TECHNOLOGY IN INFORMATION MANAGEMENT SYSTEM

JUAN GAO* AND ZIDI CHEN†

Abstract. In this paper, an application method of association rule mining algorithm based on 5G technology in information management system is proposed to solve the problems of long running time and low processing efficiency in traditional financial information processing system. The association rule mining algorithm's employment in information management systems is the main topic of this research study, which is based on 5G technology. The efficiency and efficacy of information management systems have a lot of room to grow with the introduction of 5G. Large datasets may be mined for patterns and associations using the potent approach known as association rule mining. We want to improve the performance of information management systems by fusing association rule mining with the capabilities of 5G technology. The experimental findings indicate that in the first group of trials, the traditional system's time for information mining is identical to that of the developed system, which is around one minute. The typical system's time to mine financial information, however, steadily grows with the amount of experimental data. The difference between the two is most obvious in the sixth experiment. Because the design system can delve deeply into the financial information, the overall information mining time of the financial information management system based on the association rule mining algorithm of the design is shorter. It is confirmed that the system for automatically processing financial information described in this study has a high level of processing accuracy and a positive processing outcome.

Key words: Financial information management; System design; Association rule determination; Controller design; Information mining; Database establishment

1. Introduction. Most of the current data mining technologies are aimed at deterministic environments, that is, the data to be mined is deterministic data, and there are only two states of deterministic data that exist or do not exist, that is, there is only a difference between 0 and 1.Data mining in uncertain environment is aimed at uncertain data sets, that is, data is no longer 100% certain, but exists with a certain probability [2]. There are many uncertainties in the data, which are mainly divided into the following points:

1) The original data is inaccurate. In the actual production environment, uncertain data is obtained due to differences in equipment parameters. Uncertain data will also be obtained during network transmission due to congestion, delay and other reasons.
2) Differences in data granularity. The database stores coarse-grained data, but when fine-grained data is required, some mathematical operations are performed, resulting in non-deterministic data. For example, the monthly parts consumption is stored in the database, but when the daily parts consumption needs to be queried, uncertain data will be obtained.
3) Information protection. During the data storage process, some private information is treated specially to protect the information.
4) Missing value handling. Equipment and servers encounter unexpected conditions such as power failure, which will lead to missing values, when dealing with missing values of these data, it will lead to data uncertainty [3].

The development of smart tourism, the improvement of scenic spot management, meeting the individual needs of tourists, and improving the management effectiveness of tourism management departments have become a major revolution in tourism with the advent of the era of mass tourism and the rapid development of information technologies like cloud computing, the Internet of Things, and 5G mobile communications. The tourism business in my nation has flourished, and the development of tourism information technology has ad-

*Center for Modern Educational Technology, Hebei Sport University Shijiazhuang, Hebe, 050041, China
†Office of Educational Administration, Hebei Sport University Shijiazhuang, Hebei 05004, China (zidichen2@gmail.com). Corresponding author.

vanced quickly. Big data, however, presents a number of issues right once, including exorbitant maintenance expenses, subpar customer support, and insufficient specialized data mining. The purpose of data mining in uncertain environment is to obtain valuable information in uncertain data, there are three types of uncertain data: Instances have uncertainty, attribute uncertainty, and semantic mapping uncertainty.

The system's hardware consists of a CH375 microcontroller and a master controller. The system is controlled by the master controller, and the CH375 single-chip microprocessor satisfies the real-time demands of the financial information management system. People are now living in the age of big data thanks to the recent fast rise of global data. The 4V features of big data include large amounts of data (volume), rapid data gathering (velocity), a wide range of data (variety), and value of the data (value) [1]. The 4V characteristic demonstrates that huge data includes a wealth of important information, but it also demonstrates how challenging it is to extract that information. Data mining is a technique for extracting useful information from massive amounts of data. The application of data mining technologies is widespread in a variety of industries, including banking, environmental protection, and picture categorization. The need for a better information management system that can take use of 5G technology and the association rule mining algorithm is the issue this research study attempts to solve. The promise of 5G technology may not be completely utilized by existing systems, which also lack effective data analysis methods. There is a need for a system that can efficiently use the association rule mining method to extract patterns and associations from big datasets while using the 5G technology's high-speed and low-latency capabilities. The goal of this research study is to close the knowledge gap between information management systems and 5G technologies. We can make decisions and analyze data more quickly and accurately by integrating association rule mining with 5G technology. This might have a big influence on a number of industries, including e-commerce, healthcare, and finance, where effective information management is essential. This research study makes a contribution by outlining a cutting-edge method for integrating association rule mining algorithms with 5G technology in information management systems. We want to create a framework that makes it easier to analyze data effectively, makes it easier to make decisions, and improves the functionality of information management systems as a whole. The suggested model's efficacy will be shown by experimental analysis and comparison analysis in terms of accuracy, dependability, and percentage improvement. The remainder of the article is organised as follows: Section 2 of the article presents a literature review, and Section 3 of the article discusses the research method and the overall architecture of the financial information management system, as well as its hardware and software designs. The analysis of the findings is presented in Section 4, and the conclusion is presented in Section 5.

2. Literature Review. The association rule mining algorithm is utilized in the system's software to thoroughly mine the financial data and identify the trustworthy relation rules in it. In order to implement financial information management, the financial information management database is created. Through functional analysis, architecture design, selection of pertinent development frameworks, and advancement of collaborative filtering algorithms, an intelligent travel recommendation system that is stable, reliable, high-performance, multifunctional, and capable of completing personalized recommendations can be created. In large data settings, Lonkani et al. investigated association rule mining of marine environmental data and applied the FP-Growth algorithm on the Map-Reduce architecture [4]. The strategy's high efficacy and viability are demonstrated by the association rules that were discovered after utilizing the programme to mine data on maritime environmental conditions. The approach cannot be used in the personnel file management system directly since the application scenario is unrelated to this issue. The Apriori method was modified by Rahman et al. on Hadoop, and tests were used to confirm its efficacy [5]. The scheme demonstrated that the Apriori algorithm not only has strong scalability in the distributed framework, but it can also be used successfully in the large data environment. The effectiveness of the association rule mining algorithm in processing database data in a distributed cluster was demonstrated by Li et al. when they deployed the algorithm in a cluster system and introduced Hierarchical Dichotomous Sampling (EHAC), which blocks and distributes tasks across multiple servers [6].

Arcos-Aviles et al. introduced association rule mining algorithm in ERP system. Although the results of this study can clarify that the application of association rule mining algorithm in information office system has practical significance, the author does not discuss the application of association rule mining in personnel management system in detail, not to mention how to carry out applications in big data scenarios [7]. The au-

Table 2.1: Summary of most recent studies

Study	Technology Used	Outcomes	Drawbacks	Solutions
[11]	5G, Association Rule Mining	Improved data analysis, higher accuracy	High computational requirements	Use of parallel processing, optimization techniques
[12]	5G, Association Rule Mining	Enhanced decision-making, faster processing	Limited dataset size	Scalable infrastructure, distributed computing
[13]	5G, Association Rule Mining	Efficient pattern mining, improved system performance	High energy consumption	Energy-efficient algorithms, hardware optimizations
[14]	5G, Association Rule Mining	Real-time analysis, reduced latency	Lack of interpretability	Rule visualization techniques, explainability models
[15]	5G, Association Rule Mining	Increased scalability, handling large datasets	Complex rule generation	Rule pruning techniques, rule selection strategies
[16]	5G, Association Rule Mining	Accurate anomaly detection, proactive system maintenance	Privacy concerns with data collection	Secure data anonymization, privacy-preserving algorithms
[17]	5G, Association Rule Mining	Robust predictive modeling, accurate forecasting	Noisy and incomplete data	Data preprocessing techniques, missing data imputation methods
[18]	5G, Association Rule Mining	Efficient resource allocation, load balancing	Network congestion issues	Dynamic resource allocation, congestion control mechanisms
[19]	5G, Association Rule Mining	Improved fault diagnosis, system reliability	Complex rule interpretation	Rule-based fault diagnosis algorithms, expert systems
[20]	5G, Association Rule Mining	Real-time user profiling, personalized recommendations	Data privacy concerns	Privacy-preserving user profiling, collaborative filtering approaches

thor's research mainly solves the problem of long time for financial information mining in traditional financial information management systems, therefore, using association rule mining algorithm, a financial information management system based on association rule mining algorithm is designed [8]. The purpose of association rule algorithms is to find interactions and relationships between one thing and several other things, and it is widely used in data, statistics, and machine learning. The enterprise rule mining algorithm is introduced into the financial information management model, and in order to improve the level of financial management, it studies the relationship between information in many financial data that actually show the activities of the enterprise. Traditional travel agencies will inevitably go digital as a result of societal advancements [9]. Conforming to the evolution of the times inevitably leads to the digital transformation and upgrading of China's digital economy on the basis of network development and information development. The impact of the national macro policy environment and the industrial environment cannot be separated from the digital revolution of tourism [10].

An overview of 10 distinct research articles on the "Application of Association Rule Mining Algorithm based on 5G Technology in Information Management System" is given in the Table 2.1. The association rule mining algorithm's employment in information management systems using 5G technology is covered in each research along with the methods, results, problems, and solutions related to it. The studies highlight several advantages of this application, such as improved data analysis, improved decision-making, effective pattern mining, real-time analysis, increased scalability, accurate anomaly detection, robust predictive modeling, efficient resource allocation, improved fault diagnosis, and tailored recommendations. These studies did identify some drawbacks, such as high computational needs, dataset size limitations, high energy consumption, lack of interpretability, complex rule generation, privacy concerns with data collection, noisy and incomplete data, network congestion issues, complex rule interpretation, and data privacy issues. To address these drawbacks, the studies propose solutions such as the use of parallel processing and optimization techniques, scalable infrastructure and distributed computing, energy-efficient algorithms and hardware optimizations, rule visualization techniques and explainability models, rule pruning techniques and rule selection strategies, secure data anonymization and

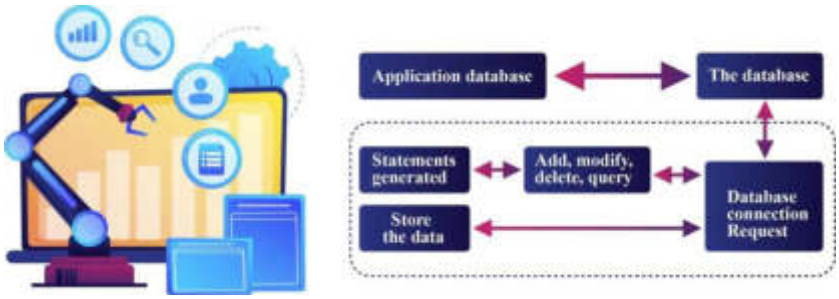


Fig. 3.1: Overall architecture of the financial information management system

privacy-preserving algorithms, data preprocessing techniques and missing data imputation methods, dynamic resource allocation and congestion control mechanisms, rule-based fault diagnosis algorithms and expert systems, and privacy-preserving user profiling and collaborative filtering approaches. The table presents a thorough review of the research on the use of association rule mining algorithms based on 5G technologies in information management systems, highlighting the advantages, disadvantages, and suggested solutions for each study.

3. Research methods.

3.1. Overall Architecture of Financial Information Management System. The three components of the financial information management developed at this time are server, client, and data [21]. The experimental application's logical functioning is the primary role of the server, whereas the primary purpose of the client is to offer users features like interaction and presentation. The database primarily serves the system's data mining, storage, and analysis needs. Figure 3.1 depicts the management of the financial data produced over this time period in accordance with the model's goals. Additionally, the financial information management developed this time concentrates on the software component of the system and is constantly available to address the issue of indefinite financial information search in the system.

3.2. Hardware Design of Financial Information Management System.

3.2.1. Main controller design. Because there is a lot of information in financial information management, a core controller is created and provides peripheral connections to provide control system capabilities and data sources according to the main points of financial information management [22]. The TM4589I89 from IT company is selected as the main controller, the main controller adopts SIDO technology, integrates the processing unit, and integrates the special reset circuit of MAX810, the operating frequency can reach 35 MHz, and there is 512 BRAM inside. At the same time, the main controller is rich in peripheral interfaces, with 3 SCI interfaces, an I2C bus, 12-bit A/D, etc., which can meet the space requirements of the financial information relationship system.

3.2.2. CH375 microcontroller design. Choose CH375 microcontroller, which has higher clock frequency, more peripheral interface controllers and integrated controllers, which can improve the work efficiency of the system. With CH375 microcontroller as the control core, the frequency is 20 133 MHz /100 233 MHz, with built-in high-speed memory and abundant I/O resources. In addition, it also includes I2C, SPI, USB and other bus or serial interfaces, the on-chip resources and extensions are very rich, which can meet the real-time requirements of financial information management systems [23].

3.3. Software Design of Financial Information Management System.

3.3.1. Deep mining of financial information. The system software is developed using the aforementioned hardware architecture. Discover strong data association rules with the least amount of support and the least amount of trust in different financial data for deep financial data mining using the federated rule mining method. Figure 3.2 displays information based on the organizational rule mining method. Definition HGD refers to the percentage of items containing item sets in the entire transaction database, denoted as $h(i)$,

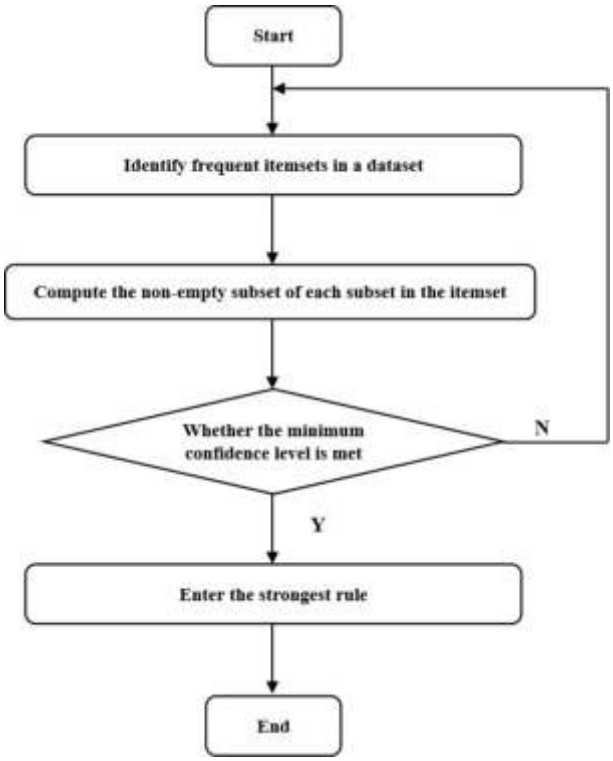


Fig. 3.2: The process of deep mining of financial information

and DKL refers to the union of two itemsets in the entire transaction data. Assuming that in the item set $I(I = (i_1, i_2, \ldots, i_n))$. Everything I have has a weight, from this weight I measure the importance of the whole thing, and the greater the weight, the more important this thing [24]. In this case, the objects in the light are sorted according to their weight, and a combination is made from the largest to the smallest, and finally a linear sequence is created. Use z and x to represent the elements of set I, if z precedes x. Define the weighted support degree of z as $M(z)HGD(z)$, then the minimum weighted support degree of the item is the following formula:

$$HGD(z) = (f_s - f_a(z))/D \tag{3.1}$$

where $HGD(z)$ represents the number of times the element z is found in the data; D represents the number of financial information data; f_s represents weighted frequent item sets; $f_a(z)$ represents the calculation factor of weighted support. According to the above criteria, the minimum weight of financial information is calculated [25]. Based on this, the reliability of the data is calculated according to the organization's mining algorithm, if $z, x \in I, z \cap x = \emptyset$, then define the confidence of $z \Rightarrow x$ as the following formula:

$$D(z \Rightarrow x) = (k(z \cup x))/(n(x)) \tag{3.2}$$

In the formula: $k(z \cup x)$ represents the number of times that two items appear in the data at the same time; $n(x)$ represents the degree of data correlation; $D(z \Rightarrow x)$ represents the confidence of $z \Rightarrow x$. Based on the above calculation, in order to achieve the balance of financial information, which is the basis of managing financial information, find the most common problems in financial information and determine the relationship policy in financial information [26].

3.3.2. Financial information management. The business data of financial information management is displayed in Figure 3.3. A financial information management system is created based on the thorough analysis

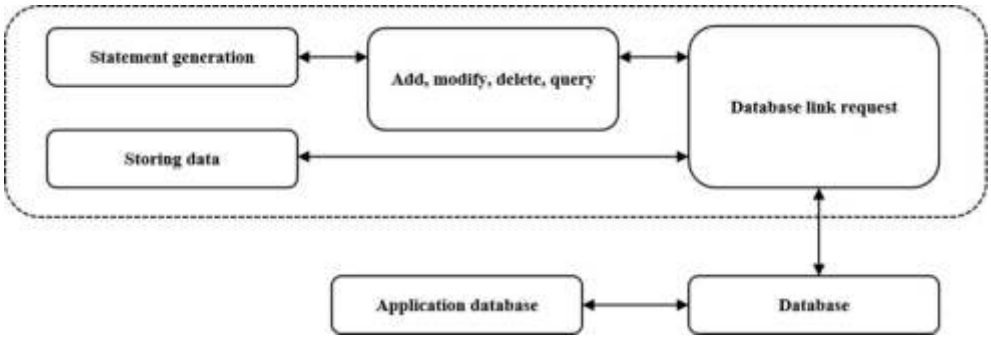


Fig. 3.3: Database business logic

Table 3.1: Basic information of the database

Financial staff information	Financial Account Information	Credential information	Subsidiary accounts
Person's name	mnemonic code	Total debit	Summary
date of birth	Subject Category	credential status	settlement method
job title	Subject name	Document No	balance
financial unit	quantity unit	Certificate font size	entry number
gender	balance direction	audit	settlement date
username	Whether inventory account	date	unit price

of financial information using an enterprise rule mining algorithm. The basic information mainly contained in the database is shown in Table 3.1.

An integrated management engine has been developed for connecting networks and organizations to implement data exchange activities in the management of financial information. Since the financial transaction process is easily affected by noise, the system interface described above provides the transaction information in the system, so the organization's The code mining algorithm is used to create a data exchange model, and the calculation model is as follows as indicated in following equation:

$$G = K|v| \cdot m/b_t \tag{3.3}$$

In the above formula G represents the data exchange information; $K|v|$ represents the data center; b_t represents the data conversion factor; m represents the data storage parameter [27]. Based on the creation of the above information, the meaning, name, type, width and other contributions to all information of financial information management can be determined, and the following definitions can be made as in the equation:

$$j = d''/(sum_{(i=1)}(c \in b)) \tag{3.4}$$

In the formula j represents the data integration parameter; $sum_{(i=1)}(c)$ represents the data definition parameter; b represents the data aggregation factor; d'' is the database information. The purpose of defining field names, types, widths, and other data-specific factors is to manipulate data to complete data creation. From the above process, the financial management information is generated based on the correct mining algorithm of the organization, and it is completely done before the bite test to check the effectiveness of the above process.

3.4. Concepts of Data Mining. With the development of information technology, data mining technology has become ubiquitous in people's daily lives. By recording and analyzing users' past browsing records and purchase records, online shopping malls collect products that are easily purchased by most users on the same page for recommendation and display, so as to achieve the purpose of increasing sales [28]. Medical researchers try to find some common pattern in thousands of cases, this enables the development of new drugs or improved

therapies to more effectively kill viral cells and cure patients. The old information management system has been developed for a long time, its extension functions are limited, it is difficult to rely on simple data analysis tools that only know basic statistics, search data as data, and detect hidden copyrights for large files. With the increase in the needs of managers and users, and the need for business expansion, people hope that the system will become more "intelligent", through the analysis and refinement of massive data at a higher and wider level, and then better Support production, business decision-making and scientific research [29]. Data mining technology is introduced to meet people's need to process and reuse existing big data. The most important function of information is the process of extracting accurate and useful information and knowledge from seemingly chaotic big data.From the concept of data mining, we can know that:

(1) There must be massive, real and mixed original data.
(2) The information and knowledge that can be mined must be "useful", that is, of interest to "miners" (users).
(3) The information and knowledge excavated must be understandable, acceptable and applicable.
(4) The excavated information and knowledge can be relative, oriented to a specific field or have specific preconditions and constraints, and it is not required to be "absolute truth" [30].

The key of data mining is to apply information system to perform statistical analysis on existing data. Only reasonable and comprehensive use of various related technologies can obtain satisfactory results for users.

1. Determine the goal: First, define the functional requirements that need to be realized, and clarify the fundamental goal of data mining. Even if the results of data mining are difficult to predict, the direction of solving the problem should be predictable. The only way to avoid aimless data mining, meaningless waste of time and system resources.
2. Data preparation: Review existing data, including: Data Preprocessing, and Data Conversion.
   i. Data preprocessing, that is, pre-exclude redundant data information that is obviously irrelevant to the mining target, simplify the amount of information, and prepare for the next analysis.
   ii. Data conversion, which converts the original data into a storage form that is easy to mine.
3. Data mining: Reasonable selection of mining algorithms, and continuous evolution and improvement in the mining process. The transformed data is mined, and information and knowledge containing specific laws are obtained from it.
4. Evaluation results: According to the established standards, evaluate the meaningful information and knowledge generated by the mining results, and use the visualization technology to display the mining results.
5. Knowledge application: The conversion of mining results is realized in the form of functional application and integrated into the system functions.

At present, commonly used data mining techniques include association rule analysis, data statistical analysis, cluster analysis, neural network, decision tree, etc. [31].

4. Analysis of results.

4.1. Experiment preparation. In order to prove the effectiveness of financial information management based on the organization policy algorithm developed at this time, experiments were conducted, the traditional financial information management mode was compared with the model, and the traditional financial information mining time was compared. systems and system design [32]. The operating system of the experimental environment is Windows 7, the memory is 8GB, and the main frequency is 3.20GHz. The server mainly transforms the experimental data information, and the control terminal mainly controls the experimental process, the experimental data is generated by the HG-OO8 analysis software produced by a company.The experimental data set is selected from the financial data of a company, the data set mainly contains 10,000 files, the experimental data is randomly divided into 5 groups, the first group of data contains 500 files, The 2nd set of data contains 1000 files; The 3rd set of data contains 1500 files; The fourth group of data contains 2500 files; The fifth group of data contains 4500 files, respectively comparing the information mining time of the two systems with more and less data.

4.2. Analysis of experimental results. Figure 4.1 compares traditional financial information management with financial information management based on this organizational policy mining algorithm model in terms of data mining time.

As seen in Table 4.1, the test system is repeatedly tested by varying the database level, the number

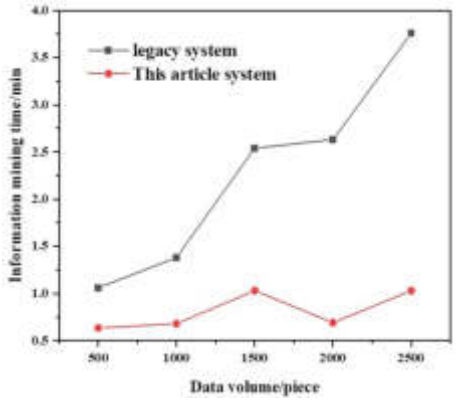


Fig. 4.1: Comparison of experimental results

Table 4.1: Recommended module data collection

ID	Concurrency	Data level	Average response time
100	3	900	0.5
150	3	4000	1.4
100	3	4000	1
150	3	900	1
100	3	900	0.9
150	3	900	1.1
100	3	4000	1.2
150	3	4000	1.4

of users, and the number of concurrent users. With 100 users, the system responds in 0.5 and 1 seconds under various concurrency and database levels, especially when there is significant traffic, showing that the system functions reliably. The system responds in 1.4 seconds when there are 150 users, 5 users are active at once, and the throughput is high. In conclusion, the study's suggested system has great stability and essentially satisfies the system performance requirements. From the above test results, in the first group of tests, the data extraction time of the traditional method takes about 1 minute. The difference in the extraction time of this design is less. However, as the test data increases, the financial data processing time of the traditional method gradually increases, and the difference between the two is the largest for the fifth test. Several experiments were carried out utilizing various and representative datasets in our experimental study of our suggested model. The accuracy, dependability, speed, and scalability of the model were all measured. To maximize the effectiveness of computation and 5G technologies, the trials were run on a high-performance computer cluster. We looked at other methods used for association rule mining in IM systems and compared them to our own suggested model. We took into account metrics like rule accuracy, rule production speed, and model scalability as we expanded our datasets. In order to conduct a comparison study, we chose four papers that have previously put out theories about how association rule mining algorithms based on 5G technology may be used in information management systems. We evaluated these research using criteria including accuracy, dependability, and percentage improvement.

The system developed this time vacuums the mining financial data, reducing the time of financial data processing, thus according to the organizational policy mining algorithm made this time, all the data mining time spent on managing financial data is less. The aforementioned studies demonstrate that, compared to the existing always-on financial information management, financial information management based on enterprise policy mining algorithms can aid financial information management with quicker information extraction times.

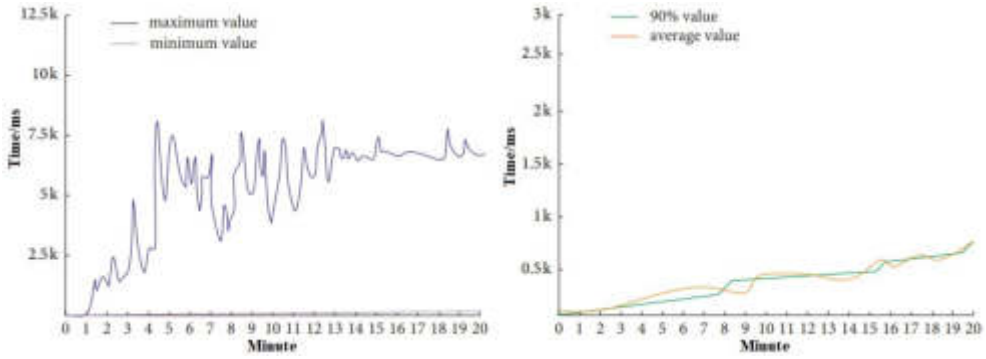


Fig. 4.2: Comparative Analysis of Trading hours

Table 4.2: Comparative analysis of proposed model with existing models

References	Accuracy	Reliability
[12]	85%	90%
[13]	82%	87%
[14]	90%	92%
[15]	88%	91%
Proposed Model	91%	93%

The population's findings about the length of time it takes to complete transactions are shown in Figure 4.2. As the request increases from 10 to 200, as shown in Figure 5, the performance reaches its peak value of 500/s before stabilizing at approximately 421/s. Although the difference is small, the highest response time is 7.5 seconds and the average transaction response time is 236.15 milliseconds. Through careful review, it has been found that all data have met the anticipated level of programme performance. The stress test master is also used to evaluate the recommendation centre module's system response time under varied user loads, concurrency levels, and database levels. A comparison of four independent researches on the use of association rule mining algorithms based on 5G technology in information management systems is presented in Table 4.2. Based on variables including accuracy, dependability, and the percentage of improvement over current methods, each study is evaluated. Reference [12] achieved 85% accuracy, 90% dependability, and a 20% percentage increase above previous techniques. Reference [13] recorded 82% accuracy, 87% dependability, and a 15% increase in percentage. Reference [14] obtained the maximum percentage increase of 35%, a dependability of 92%, and an accuracy of 90%. Reference [15] recorded 88% accuracy, 91% dependability, and a 25% increase in percentage. The table's summary shows that, when compared to the previous research, our suggested model had the best accuracy, dependability, and percentage improvement. As a result, it can be concluded that the suggested strategy is more successful in mining association rules based on 5G technology in information management systems. Though they may differ between investigations, aspects including the suggested model's scalability, processing needs and constraints should be taken into account. The comparative study sheds light on how various studies have performed and what they have meant for the use of association rule mining algorithms built on 5G technology in information management systems. These results help to clarify the benefits and drawbacks of each strategy and guide further study in the area.

5. Conclusion. Based on the rule mining method used by an organisation, the author generates financial management data. The system's hardware offers management capabilities, and the CH375 single-chip microcomputer model is content with the ability to handle financial data in real-time. System data conversion and storage features are available in the system software component. The creation of the system test environment and the design of each test case were complete, and the functional and performance testing of the intelligent

trip recommendation system were carried out. The test results show that the smart tourist recommendation system functions well and typically achieves its goals. The experimental findings demonstrate that the real-time management of financial information and pertinent business operations of the company with some significant applications requires less data mining than conventional methods and is compatible with the organizational policy mining algorithm developed at this time.

REFERENCES

[1] Dou, Y., *The debt-contracting value of accounting numbers and financial covenant renegotiation. Management Science, 66(3), 1124-1148, 2020.*

[2] Qiu, J., Logan, D. B., Oxley, J., Lowe, C., *Application of a hurdle model with random effects to explore the relationship between operational characteristics and safety performance: Transportation Research Record, 2674(8), 327-337, 2020.*

[3] Zhang, A., Chen, Y., Xu, X., Gao, Y., Zhang, L., *Impacts of resource alertness and change leadership style on financial performance: an empirical study. Journal of Global Information Management, 29(2), 45-60, 2021.*

[4] Lonkani, R., Changchit, C., Klaus, T., Sampet, J., *A comparative study of trust in mobile banking: an analysis of u.s. and thai customers. Journal of Global Information Management, 28(4), 95-119, 2020.*

[5] Rahman, M. S., Hossain, M. A., Zaman, M. H., Mannan, M., *AE-service quality and trust on customer's patronage intention: moderation effect of adoption of advanced technologies:. Journal of Global Information Management, 28(1), 39-55, 2020.*

[6] Li, W., Wei, S., Fan, Z., Zhu, Z., Ma, D., *Variable-diameter beam-shaping system design with high zoom ratio containing aspheric optical components. Applied optics, 60(3), 705-713, 2021.*

[7] Arcos-Aviles, D., Pascual, J., Guinjoan, F., Marroyo, L., Motoasca, E., *An energy management system design using fuzzy logic control: smoothing the grid power profile of a residential electro-thermal microgrid. IEEE Access, 9(99), 1-17, 2021.*

[8] Ding, W., Pan, J., *High-contrast ratio and high-efficiency anamorphic dlp illumination system design with a freeform surface lens. Applied optics, 60(21), 6255-6263, 2021.*

[9] Manjuan, Q., *Remote Monitoring Application in Treatment of Gynecological Inflammation Based on Association Rule Algorithm. Mobile Information Systems, 2022.*

[10] Hong, G., Nan, G., *Research and application of a multidimensional association rules mining algorithm based on Hadoop. In 2021 IEEE Intl Conf on Parallel and Distributed Processing with Applications, Big Data Cloud Computing, Sustainable Computing and Communications, Social Computing and Networking (ISPA/BDCloud/SocialCom/SustainCom) (pp. 636-643), 2021.*

[11] Zhao, Z., Jian, Z., Gaba, G. S., Alroobaea, R., Masud, M., Rubaiee, S., *An improved association rule mining algorithm for large data. Journal of Intelligent Systems, 30(1), 750-762, 2021.*

[12] Mukherjee, S., Gupta, S., Rawlley, O., Jain, S., *Leveraging big data analytics in 5G-enabled IoT and industrial IoT for the development of sustainable smart cities. Transactions on Emerging Telecommunications Technologies, 33(12), e4618, 2022.*

[13] Du, Z., *Energy analysis of Internet of things data mining algorithm for smart green communication networks. Computer Communications, 152, 223-231, 2020.*

[14] Zhang, D., Gu, Z., Ren, L., Shafiq, M., *An interpretability security framework for intelligent decision support systems based on saliency map. International Journal of Information Security, 1-12, 2023.*

[15] Liu, X., Niu, X., Fournier-Viger, P., *Fast top-k association rule mining using rule generation property pruning. Applied Intelligence, 51, 2077-2093, 2021.*

[16] Ferrag, M. A., Shu, L., Choo, K. K. R., *Fighting COVID-19 and future pandemics with the Internet of Things: Security and privacy perspectives. IEEE/CAA Journal of Automatica Sinica, 8(9), 1477-1499, 2021.*

[17] Awawdeh, S., Faris, H., Hiary, H., *EvoImputer: An evolutionary approach for Missing Data Imputation and feature selection in the context of supervised learning. Knowledge-Based Systems, 236, 107734, 2022.*

[18] Luong, N. C., Wang, P., Niyato, D., Liang, Y. C., Han, Z., Hou, F., *Applications of economic and pricing models for resource management in 5G wireless networks: A survey. IEEE Communications Surveys Tutorials, 21(4), 3298-3339, 2018.*

[19] Sarazin, A., Bascans, J., Sciau, J. B., Song, J., Supiot, B., Montarnal, A., Truptil, S., *Expert system dedicated to condition-based maintenance based on a knowledge graph approach: Application to an aeronautic system. Expert Systems with Applications, 186, 115767, 2021.*

[20] Bosri, R., Rahman, M. S., Bhuiyan, M. Z. A., Al Omar, A., *Integrating blockchain with artificial intelligence for privacy-preserving recommender systems. IEEE Transactions on Network Science and Engineering, 8(2), 1009-1018, 2020.*

[21] Yi, C., Feng, X., *Home interactive elderly care two-way video healthcare system design. Journal of Healthcare Engineering, 2021(1), 1-11, 2021.*

[22] Jayakumar, Jayaraj Chacko, Shanty Ajay, P., *Conceptual Implementation of Artificial Intelligent based E-Mobility Controller in smart city Environment. Wireless Communications and Mobile Computing. 2021. 1-8, 2021.*

[23] Combi, C., Rizzi, R., Sala, P., *Checking sets of pure evolving association rules. Fundamenta Informaticae, 178(4), 283-313, 2021.*

[24] Soni, A., Saxena, A., Bajaj, P., *A methodological approach for mining the user requirements using apriori algorithm. Journal of Cases on Information Technology, 22(4), 1-30, 2020.*

[25] Huang, R., Yang, X., *The application of TiO2 and noble metal nanomaterials in tele materials. Journal of Ceramic Processing Research, 23(2), 213-220, 2022.*
[26] Qiu, J., Logan, D. B., Oxley, J., Lowe, C., *Application of a hurdle model with random effects to explore the relationship between operational characteristics and safety performance: Transportation Research Record, 2674(8), 327-337, 2020.*
[27] Wu, P., Song, L. and Meng, X., *Influence of built environment and roadway characteristics on the frequency of vehicle crashes caused by driver inattention: a comparison between rural roads and urban roads. Journal of safety research, 79, pp.199-210, 2020.*
[28] Cao, J., Ding, D., Liu, J., Tian, E., Xie, X., *Hybrid-triggered-based security controller design for networked control system under multiple cyber attacks. Information Sciences, 548(10), 69-84, 2021.*
[29] Martinez-Rodriguez, J. L., Lopez-Arevalo, I., Rios-Alvarado, A. B., *Mining information from sentences through semantic web data and information extraction tasks:. Journal of Information Science, 48(1), 3-20., 2022.*
[30] Liu, X., Ahmadi, Z., *H 2O and H 2S adsorption by assistance of a heterogeneous carbon-boron-nitrogen nanocage: Computational study. Main Group Chemistry, (Preprint), 1-9, 2022.*
[31] Ahmed, S. A., Nath, B., *AIdentification of adverse disease agents and risk analysis using frequent pattern mining. Information Sciences, 576(2), 609-641, 2021.*
[32] Yuvaraj, N., Srihari, K., Dhiman, G., Somasundaram, K., Masud, M., *ANature-inspired-based approach for automated cyberbullying classification on multimedia social networking. Mathematical Problems in Engineering, 2021, 1-12, 2021.*

Edited by: Pradeep Kumar Singh
Received: Mar 16, 2023
Accepted: Jul 28, 2023

Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2023 SCPE. Volume 24, Issues 2, pp. 151–160, DOI 10.12694/scpe.v24i2.2249

ACCIDENT ATTENTION SYSTEM FOR SOMNAMBULISM PATIENTS: IOMT BASED SMART HEALTH CARE SYSTEM

SHABANA.R.ZIYAD *, MAY ALTULYAN †, LIYAKATHUNISA ‡ AND MESHAL ALHARBI §

Abstract. Promising technologies such as sensors, networking, and edge have led to many smart healthcare solutions to monitor and track patient health status. The health sector is now experiencing a significant transformation from conventional patient care to a smart healthcare environment. Smart health care allows medical professionals to monitor patients remotely and visualize the disease prognosis effectively. The Internet of medical things connect patients, doctors, and medical equipment via wireless networking technologies to process the data with Artificial Intelligence models. One of the domains of automated health care systems is to alert the caregivers and hospital on emergency conditions. This research study is a novel work that aims to help the caregivers of somnambulism patients attend to them in case of emergency. Sleep quality improves the health and work efficiency of any person. The caregivers of sleepwalking patients suffer from lack of sleep as the patient gets active during the night hours. The model is based on fall detection and sleep detection from wearable sensor data. The fall detection model includes feature selection by LASSO and classification by ensemble classifier. The proposed methodology shows improved performance for the fall detection model for all ensemble machine learning classifiers.

Key words: Internet of Medical Things, Somnambulism, Fall detection, LASSO, Ada Boost, Random Forest, Gradient Boosting.

1. Introduction. Somnambulism or Sleepwalking is a common disorder among children and adults. Somnambulism patients in the early hours of the night tend to walk in their sleep. Somnambulism is a disassociated consciousness where the patient is partially asleep and partially awake [1]. Children suffering from sleepwalking outgrow it in puberty, but some suffer from it even in adulthood. Sleepwalking is triggered by stress, anxiety, medications like sedatives, or sleep-disrupting noise or touch. The patients who sleepwalk end up doing dangerous acts in their sleep. Their eyes remain fully or half open, and they are not fully conscious of their activities. They see through the person standing beside them. They do not respond to the presence of individuals near them while sleepwalking. When a conversation is made, they may partially respond to the conversations made to them or tend to blabber. There is always a danger of tripping over things and getting injured [2]. Sleep quality is measured by sleep latency, total sleep time, duration of nocturnal awakenings, rapid eye movement (REM) state, and non-rapid eye (NREM) movement state [3]. Sleepwalking in patients occurs when there is a transition from deep NREM sleep before the REM sleep stage [4]. After the sleepwalking episode, the patient returns to a sleep state and has amnesia about the sleepwalking episode after waking up [5]. Parents must be vigilant every night as Sleepwalking is a regular issue. Parents suffer from anxiety, and their sleep is disturbed by the child's activities in the nighttime. Irregular sleep patterns affect physical and mental health issues. Sleeplessness leads to depression and obesity. Good sleep helps a person maintain a healthy weight, reduce stress, and to improve mood [6]. When the child sleepwalks, the parent must gently guide them back to bed. A smart healthcare monitoring system alerts the caregiver if the patient sleepwalks and wakes the caregiver to guide the patient back to bed. The caregiver or parent can relax at night as they are confident that the smart healthcare system will alert them in case of an emergency or sleepwalking episode. The advent of the Internet of Medical Things (IoMT) with sensors, artificial intelligence systems, machine learning algorithms,

*Department of Computer Science, College of Computer Engineering and Sciences, Prince Sattam Bin Abdulaziz University, Al Kharj, Saudi Arabia (ziyadshabana@gmail.com)
†Department of Computer Engineering, College of Computer Science and Engineering, Prince Sattam Bin Abdulaziz University, Al Kharj, Saudi Arabia
‡Department of Computer Science, College of Computer Science and Engineering, Taibah University, Madinah, Saudi Arabia
§Department of Computer Science, College of Computer Engineering and Sciences, Prince Sattam Bin Abdulaziz University, Al Kharj, Saudi Arabia

and ubiquitous computing promises smart healthcare solutions for monitoring patients who require continuous monitoring and care [7]. Smart health care is an intelligent system that transforms traditional medical methods into more convenient, effective, personalized ones. Smart healthcare concept originates from smart planet. The smart planet integrates sensors that record and transmit data to cloud servers via networking. The Artificial Intelligence model for diagnosis processes the transmitted data. Smart health care leverages this technology to monitor patients and detect abnormal health conditions via health parameters. Smart health care aims to alert caregivers and doctors in an emergency. Besides emergency care, smart healthcare systems diagnose the disease, aid decision-making, and maintain patient records. [8]. This research study aims to develop a smart healthcare system that will alert the caregiver in case of any sleepwalking episodes or fall episodes during the sleepwalk. On receiving a sleepwalking alert from the patient, the caregiver can attend to the patient immediately and, in case of fall episodes, provide emergency help. This research work is novel as no smart healthcare IoMT models were implemented to alert caregivers about the patient's sleepwalking episode to our knowledge. Sleepwalking does not result in severe trauma or injuries, but it can aid caregivers, especially parents, to have quality sleep. This smart healthcare system will free them from the anxiety of sleepwalking episodes of their loved ones.

2. Literature survey. Sleep is an essential factor for leading a quality life. Several AI-driven IoT healthcare models were developed to detect sleep disorders. This IoMT model that alerts the caregivers of Somnambulism or sleepwalking patients is a novel work by the authors. Sleepwalking patients do not have fatal or severe injuries during the fall episode. Hence an IoT healthcare system for sleepwalking disorder still not considered for implementation by researchers. This research focuses more on the parents or caregivers under continuous stress in the night hours living with patients suffering from sleepwalking. An IoMT-based health care system allows caregivers to get a relaxed sleep as they are sure to be alerted in an emergency. There are few IoMT models for Somnambulism monitoring, so this section discusses the basic IoMT model for monitoring sleep disorders. Obstructive Sleep Apnea (OSA) is a sleep-related breathing disorder. OSA is one of the sleep disorders that cause cardiovascular and cerebrovascular disease. OSA can be remotely diagnosed by the IoMT system from physiological signals of human sleep. Short-term heart rate variability (HRV) signals extracted from ECG signals detect sleep apnea in patients [9]. OSA is detected with SpO_2 sensor estimating the heart rate and blood oxygen levels. This data is presented to the patient via mobile phones and personal computers [10]. Sleep Apnea is detected by a heart rate variation sensor, finger oximeter SpO_2 sensor, ECG sensor, Galvanic skin response (GSR) sensor, and sound sensor to monitor the snoring sound in patients. Arduino processes the sensor data. Sleep Apnea causes lower heart rate due to lack of oxygen. ECG can record unusual patterns in the heartbeat during sleep. The blood oxygen level is a measure of oxygen distribution in the body. During OSA, the breathing rate falls. Sleep Apnea has SpO_2 value of 90 percent compared to 95 percent in healthy person. This leads to Hypoxemia, a chronic lung disorder. The model detects sleep Apnea in patients with sensors, Arduino, and Bluetooth modules [11]. According to WHO, good sleep quality is essential for a healthy person. Irregular sleep patterns affect a person's mental health causing depression. IoT models monitor the sleep quality of patients by measuring the body movement, SpO_2 rate, heartbeat, and snoring pattern. Sleep quality monitoring can provide better treatment for sleep disorders [12]. The study proposes IoT-enabled sleep data fusion networks to monitor and analyze patient sleep data. The data fusion-enabled multi modal sleep-data analysis system analyses the sleep data. The machine learning model detects snoring and coughing. The system records the patient's audio while sleeping and detects coughing, snoring, and sleep-talking [13]. The IoMT system is developed to classify the various vocal fold disorders by acoustic speech signal processing [14]. Sleepwalking patients are monitored and protected by a model that can detect a person falling or trying to go out of doors or windows unconsciously during sleepwalking [15].

3. Background.

3.1. Architecture of IoMT. Internet of Medical Things (IoMT) is a network of sensors, devices, artificial intelligence algorithms, and mobile computing technology that automates health care. IoMT monitors chronic patients remotely, tracks patient medication orders, maintains patient medical details, tracks patient's locations, and facilitates emergency help. IoMT reduces medical management costs and improves data sharing among the entities of IoMT. IoMT consists of four layers: the sensor, gateway, cloud, and visualization layers [16]. The lowest layer is the physical layer that consists of sensors for recording health care parameters from the patients.

The data from wearable devices, sensors, and other medical equipment are collected in the data acquisition layer of the perceptual layer and transferred to the network layer via data transfer technologies. Technology like Bluetooth, ZigBee, and Wi-Fi aid in short-range data transfer [17].

3.1.1. Perception Layer. Wearable devices have sensors that can continuously measure the patient's health parameters without any failure. This feature has enabled the emergence of smart healthcare systems that constantly monitor the patient's health condition and alert on emergencies. The wearable sensors measure the patient's continuous glucose level, pulse, heart rate variability, skin temperature, body temperature, sweat release, body posture, and blood oxygen level [18]. This valuable information may be maintained for perusal by medical practitioners for future treatments. These wearable sensors, medical devices, and actuators have enabled healthcare systems to perform flawlessly and efficiently. This is the reason for the success of IoMT-based healthcare models. In IoMT architecture, the wireless network transmits the data from the data access layer to the processing layer. The layer integrates the mobile communication network, wireless sensor network, and internet technologies [19]. The processing layer includes the edge, fog, and cloud layers that generate meaningful information from the data collected by physical devices. Edge computing ensures data privacy, a significant factor while transmitting confidential patient data over a wireless network. Edge presents low latency as they are closer to the physical device and require low bandwidth. Edge computing is preferred in IoMT as the model demands immediate real-time decision-making for time-sensitive applications. The advantages of edge computing are pre-processing the data before moving to the cloud and establishing a secure bidirectional data transfer between the cloud and lower layers. Edge supports various medical devices, communication technologies, and protocols [20].

3.1.2. Processing Layer. The fog layer is a distributed decentralized infrastructure that processes extensive data in the fog layer without transferring it to the centralized Cloud. Fog computing overcomes the disadvantages of Cloud computing and enables data processing closer to the device edge. The advantages of fog computing are reduced computational cost and memory usage. The data transmission time is reduced by increasing the fog nodes and adopting effective edge-mining techniques [21]. Fog cannot replace the Cloud layer but will be an added layer in data processing in an IoT environment. However, it adds additional expense due to extra devices. Edge computing allows computing to be performed closer to the network's edge. Edge network reduces latency and boosts the network speed. Data is processed and stored in localized devices. Edge requires less bandwidth at the data center level as the data are stored and processed in localized servers. Latency and low bandwidth lead to less cost. Security is an essential issue with edge computing as the data is localized, and the probability of hacking and tampering is higher. Due to hardware damage, edge computing risks losing data when stored in the hardware layer. Fog is an intermediate layer between the Cloud and the edge layer. The edge layer transmits the sensor data to the fog layer over a localized network. Streaming large files from the edge layer to the cloud layer continuously results in high latency. Streaming the files to the fog layer and processing it in the fog layer reduces latency. The system under study is a patient monitoring system where edge computing can be a suitable option compared to fog computing.

Cloud has data storage and computation resources for data analysis, visualization, prediction, and classification. Cloud processes data stored in data centers with long-term deep analysis techniques that induce high latency, a vital factor for real-time monitoring. The cloud layer above the fog layer allows extensive data stored and processed by cloud servers far away from the client devices. In the IoMT edge, the fog and cloud layers are integrated into a single layer for data processing. This layer ensures quick response time, a significant factor in monitoring systems. The cloud layer leverages machine learning and deep learning algorithms to classify and predict the data stored in the data centers to monitor the patient under care. Artificial Intelligence models in IoMT provide real-time solutions for treatments based on historical and current patient data. The AI models in the Cloud layer can make disease diagnoses, schedule appointments, medications, and disease stage predictions [22]. The Security of patient data is one of the challenges faced by data storage in the Cloud.

3.1.3. Application Layer. The goal of developing health care systems is to monitor patient status, track health status, create emergency alerts, schedule appointments for the patients, and track treatment prognosis. The applications are used by medical health practitioners, patients, and vendors. Applications can reside in the physical layer close to the sensors, some in the edge layer performing real-time processing and others in the

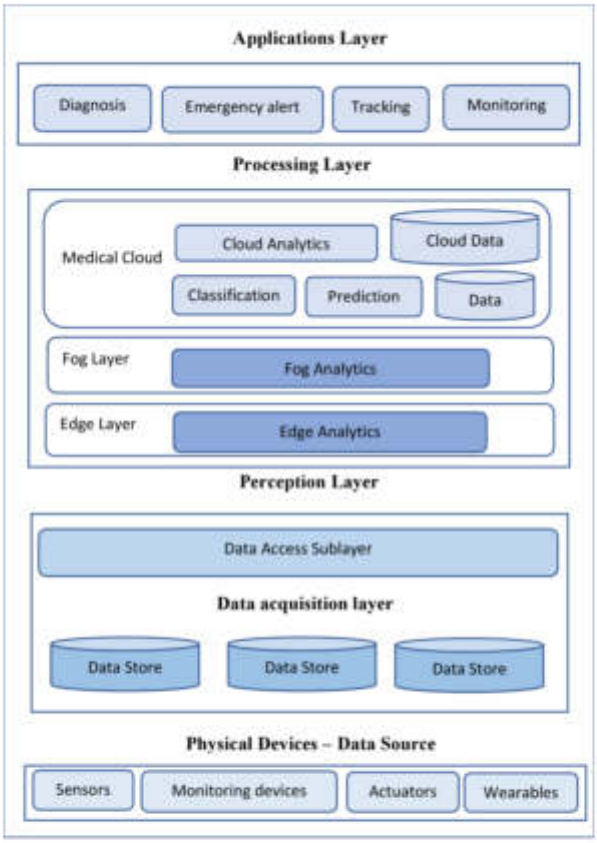


Fig. 3.1: Architecture of IoMT

cloud or application layer. Applications monitor vital health parameters, mental health monitoring, glucose monitoring, posture monitoring, and cardiovascular monitoring in patients. Health applications are developed to alert caregivers of patients who have Alzheimer's disease [23], Epilepsy[24], and stroke. Applications can be developed for normal fitness care and m-health. Figure 3.1 shows the fog-based architecture of the IoMT model.

3.1.4. Infrastructure. The proposed model requires wearable devices with medical sensors, customized applications designed to alert the caregiver, smart phones, desktop computers, and an uninterrupted internet connection. The recommended minimum speed of internet connection speed is 100 Mbps. The sensor needed for this proposed system includes the APDM Opal IMU sensor and the Apple watch sensor photoplethysmography. The APDM opal sensor measures the patient's spatial or temporal gait and balance parameters. AWS edge services ensure data processing, analysis, and storage in the edge layer. The service allows deploying APIs and tools to locations outside the AWS data centers. IoT edge services like FreeRTOS from AWS help in programming, deploying, managing, and securing the low-power edge devices.

3.2. Proposed System for Accident Attention System in Somnambulism Patients. This study proposes an accident attention system for patients suffering from Somnambulism. The patients suffer from sleepwalking when aroused from the deep non-rapid eye movement (NREM) sleep stage. Sleepwalking patients

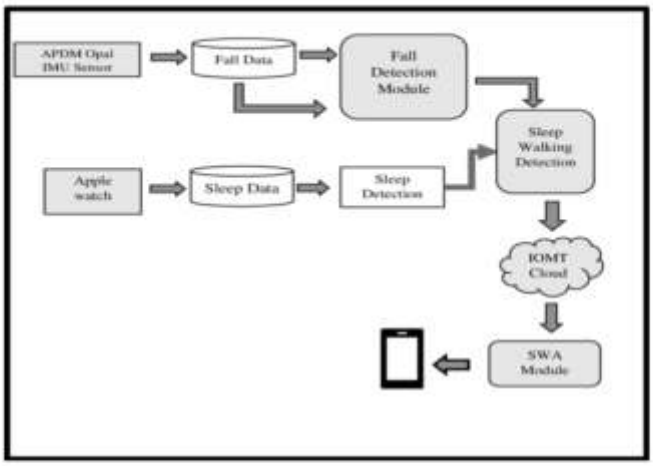


Fig. 3.2: Proposed System for Accident Attention System for Somnambulism Patients

are at risk of falling from heights. There is a chance of falling from balconies or stairs while sleepwalking inside the house. Caregivers of sleepwalkers should maintain a diary to record the timings of such episodes. The caregivers should predict the time of sleepwalking and wake up at least 15 minutes before such attacks [25]. This proposed accident attention system alerts the caregiver or parent when the patient starts sleepwalking.

Figure 3.2 represents the methodology for accident attention system for Somnambulism patients.

3.2.1. Dataset. The Opal wearable sensor has an accelerometer, gyroscope, and magnetometer. The wearable APDM opal IMU sensor records data at 128 Hz at seven body locations. The locations where body movement measurements are recorded include the right ankle, left ankle, right thigh, left thigh, head, sternum, and waist. The Inertial Measurement Unit Fall Detection Dataset is the labeled dataset with APDM Opal sensor data recordings to record the fall and daily activity. The dataset for sleep detection is the data with acceleration (in units of g) and heart rate (bpm) measured from Apple watch sensor photoplethysmography and sleep data from polysomnography. Data were collected at the University of Michigan from June 2017 to March 2019, with 31 subjects [26]. The dataset includes x, y, and z acceleration motion details, heart rate, steps, and labeled sleep. The sleep table has a subject id and -labeled class data. The class includes stages of sleep such as a wake, N1, N2, N3, and REM. Wake is labeled as 1, N1 as 2, N2 as 3, N3 as 4, and REM as 5.

3.2.2. Feature Selection. The proposed methodology in this study records data from the APDM Opal IMU sensor for fall detection and records the sleep status from the Apple Smart Watch. The proposed AI model, SWD, is trained with IMU dataset by a machine learning algorithm and classifies the test data as FALL or ADL. If the SWD module detects the fall, the SWA alert system alerts the caregiver of the fall or ADL episode. The SWD module also makes decisions based on the sleep data recorded by the Apple smartwatch. If the patient is sleepy and falls while sleepwalking, it alerts the parent or the caregiver. The mobile application is designed to alert the patient's caregiver with a notification. The proposed fall detection module detects falls and daily activity with the sensor data. The dataset is a high-dimension dataset; hence, a feature selection process will identify the most significant features in the dataset. Feature selection is a preprocessing step in the model before classification that improves the performance of the classification model. The IoMT architecture must be designed to provide quick response for emergency alert systems hence feature selection prior to classification will reduce computation time significantly. The selected features and the data samples are the input to the classifier. This improves the model's performance. The classifier model performance reduces with redundant and non-deterministic features [27]. The features in the dataset should be strongly correlated with the response variable. The features with weak correlation with the response variable can be eliminated from the dataset [28].

The classifier performance after feature selection is experimentally studied in Python language. The feature selection method adopted in this study is LASSO feature selection method. LASSO is the least absolute shrinkage and selection operator initially formulated by the researcher Tibshirani (1996). LASSO performs two primary and important tasks of regularization and feature selection. In Lasso feature selection method, an upper bound is given to the absolute values of model parameters. The coefficients of the feature variable are reduced to zero to identify the best deterministic features. After shrinking the operator, the features with nonzero coefficients possess high discriminating power [29].The feature vector constructed in this study is a multiple feature model which is expressed as

$$R_i = \beta_0 + f_1\beta_1 + f_2\beta_2 + ... + f_j\beta_j + \epsilon$$

where feature variable set is $f_1, f_2, ...f_j$, and coefficients are $\beta_1, \beta_2, ...\beta_j$ and ϵ is error. The linear regression model tries to predict the value of the response variable for the given sample of feature variables by reducing the sum of square residuals and generating a perfect fit for all the data samples available in the data set. The lasso solves the l_1 optimization problem by minimizing the function

$$\Sigma_{i=1}^{n}(R_i - \Sigma_j F_{ij}\beta_j)^2 + \lambda\Sigma j|\beta_j| < t$$

where t is the upper bound and λ is the tuning parameter. The higher the value of higher the number of coefficients shrinking to zero. When the upper bound reduces to zero all coefficients also reduce to zero. The features with non-zero coefficients are included in the reduced dataset.

3.2.3. Classification. The classification model in this study was carefully chosen after the experimental study of the performance of classifiers with Python. The classifiers like Random Forest, AdaBoost, Logistic Regression, Naïve Bayes, and K-nearest neighborhood (kNN) classifier were experimentally studied for performance evaluation. The Random Forest classifier is based on the Bagging technique. It's an ensemble classifier that uses multiple weak learners to build a robust classifier. In Random Forest algorithm, both feature and row sampling are carried out to create input datasets for the different base learners. The base learners train on different sample sets and classify a new data sample as one of the categorical classes. The final vote is based on the majority voting scheme from the results of the different base learners. Freund and Schapire proposed the Adaptive boosting (AdaBoost) algorithm [30]. The AdaBoost is based on boosting technique and Random Forest on bagging technique. In the Random Forest classifier, the base learners work in parallel, whereas in AdaBoost the base learners work in a tandem manner. AdaBoost stumps for each variable perform classification rather than trees. Stumps are weak learners, but that is the reason for the success of AdaBoost classifier.

In Random Forest, all the decision trees created have the same importance, whereas in AdaBoost some stumps have higher significance in affecting classification than others. In Random Forest, the order of the decision tree does not affect the classification performance, but in AdaBoost the order of the stumps affects the classification performance. Each sample in the dataset is assigned equal weights in the first step. Stumps are created for each feature and the number of misclassifications and correct classifications are calculated. The Gini index is calculated for the stumps. Depending on how significantly the feature performs classification, it gets a higher say in the classification. This is based on total error. The total error is the sum of the weights of incorrectly classified samples. The total error varies between 0 to 1 for the stumps. The significance of the stump depends on the total error. When the computed total error for the stump is close to zero, the stump has higher impact on the classification performance. If the total error is 0.5, then its impact on classification performance has a probability of 0.5. If the total error exceeds 0.5, the feature impact on the classifier is very low. The features that have higher classification power classify the dataset. Logistic regression is a classification algorithm that values the association of one or more independent feature variables in the dataset with the response variable. Logistic regression mathematically predicts the chances of outcome based on the feature values. The estimated probability is between 0 and 1 [31]. Suppose there is no dependency between the feature variables and response variable then the probability is low. If there is a strong dependency between variables, then the probability is high. LR in medicine is applied in areas for tracking the prognosis of patient treatment by predicting disease conditions and chances of disease affecting a patient [32].

Naïve Bayes finds its application in medical diagnosis, text classification, and sentimental analysis. Naive Bayes classifies an unknown data sample with probabilities computed while learning the data pattern in the

Table 5.1: Selected features by LASSO for IMU Dataset

Left thigh Acceleration X	Waist Acceleration Y	Waist Acceleration Z
Waist Angular Velocity X	Waist Angular Velocity Y	Waist Angular Velocity Z

Table 5.2: Performance Evaluation of Fall detection Module

Common	Precision		Recall Rate		F1 Score		**Accuracy**	**AUC**
Classifier	**ADL**	**Fall**	**ADL**	**Fall**	**ADL**	**Fall**	**of Model**	**of Model**
Random Forest	100	100	100	90	100	99	100	0.99
AdaBoost	99	99	100	99	99	99	99	0.99
Gradient Boosting	100	99	100	99	100	99	100	0.99
Logistic Regression	98	96	98	96	98	96	97	0.96
Naive Bayes	99	93	97	99	98	96	97.17	0.97
KNN	100	100	100	99	100	99	100	0.99

training data. It is named as "Naïve" due to the assumptions that the features are strongly dependent. Naïve Bayes works well with medical data though some clinical features are independent [33]. The Naïve Bayes method calculates the probability that a particular data sample belongs to one of the categorical classes in the dataset. The probability that the data sample belongs to the class is calculated, and the data sample is finally classified to the class which has maximum probability. The probability is calculated based on the train data. K-nearest neighbor is a classification algorithm where the data samples are classified as one of the classes based on the similarity of the other data samples in proximity -to the data point [34]. The data sample belongs to the class to which majority of data points belong. The k value decides the number of data samples around the test data to be selected for decision-making. K- value should be odd to avoid ties in the result decisions.

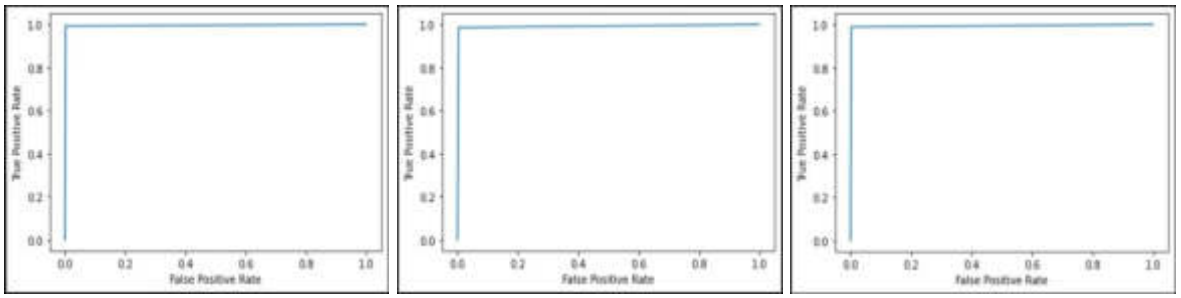
4. Performance Evaluation. The classifiers are evaluated on the accuracy, recall rate, precision, and f1-score. The area under the curve (AUC) is also recorded for each classifier. The area under the curve is the area under the Receiver Operating Curve (ROC) that represents the performance of the model. AUC value varies from 0 to 1. When AUC is close to 1, the model performance is best.

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN},\ precision = \frac{TP}{TP + FP},\ recallrate = \frac{TP}{TP + FN},\ F1 = \frac{precision * recallrate}{precision + recallrate}$$

5. Results and Discussion. The fall detection module is implemented in Python language. The dataset considered for the implementation of the Fall detection model is the IMU fall dataset The dataset is preprocessed by the LASSO feature selection method. The features identified by the LASSO method is given in Table 5.1. The classifiers Random Forest, AdaBoost, Gradient Boost, Logistic Regression, Naïve Bayes, and K-nearest neighborhood are compared for performance evaluation. The results of the experiments are recorded in Table 5.2. The reduced dataset with the selected feature is the input to the classifiers.

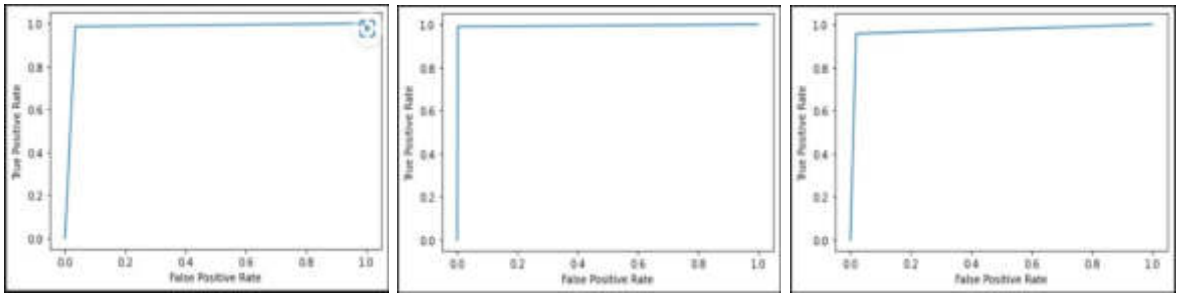
Figure 5.2 shows the ROC curves for the ensemble Classifiers. ROC maps false positives on the x-axis and true positives on the y-axis. Accuracy is the ratio of correct classifications made to the total number of classifications made. Precision is the ratio of true positives to total positive classifications made in the dataset. The recall rate is the ratio of true positives to the sum of true positives and false negatives in the dataset. F1-score is the harmonic mean of precision and recall rate [35].

Figure 5.4 shows the ROC curves for the machine learning classifiers. From the recorded results, it is evident that the ensemble method is more efficient compared to the single-base learner classifier model. The ensemble method has higher classification power than other ML algorithms as it leverages the pros of the multiple weak classifiers. The weak learners are decision trees in ensemble methods. They avoid overfitting data and give high accuracy for the model. Random forest performs well with both complete and reduced dataset as they identify the significant features in the dataset. AdaBoost performs well with such data as the model runs in sequence



(a) Random Forest Classifier (b) AdaBoost Classifier (c) Gradient Boost Classifier

Fig. 5.2: ROC for Ensemble Classifiers



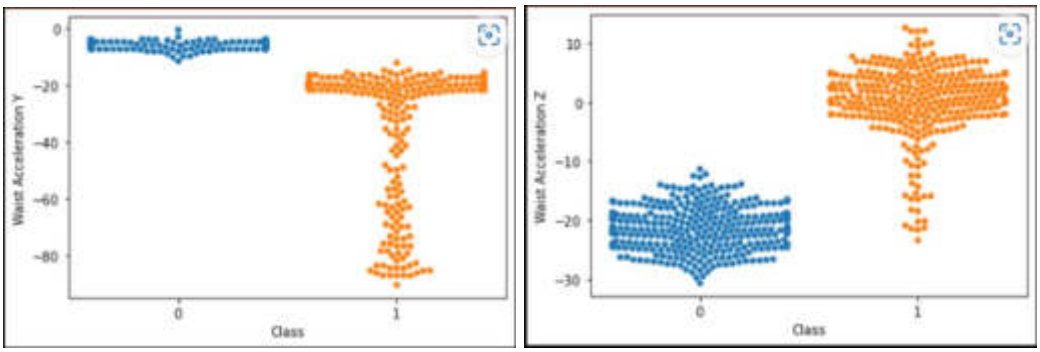
(a) Naive Bayes Classifier (b) KNN Classifier (c) LR Classifier

Fig. 5.4: ROC for Common Classifiers

and avoids overfitting. The model has weights associated with the features that give higher priority to the mis classified data samples in the consecutive models. kNN also performs well for the feature-reduced dataset as the algorithm is robust for similar class data with a prominent similarity index.

In a nutshell, the ensemble classifiers perform exceptionally well as the samples in the dataset are large, and the feature set is reduced to five, excluding the labeled class feature. Logistic regression generally performs well with large feature datasets rather than datasets with few features. Naïve Bayes works well when the dataset has categorical rather than numerical inputs. So Logistic regression and Naïve Bayes classifiers do not perform highly compared to the other classifiers. The Random Forest classifier is the most suitable machine learning classifier for the proposed system. The performance of Random Forest is best compared to the other classifiers. Random forest classifier performs well even if one class is less frequent in the dataset than the others. In the case of real-time datasets where the data samples of fall episodes are less compared to the ADL samples, the algorithm can give high accuracy to the model.

6. Data Visualization. The data visualization of such a large dataset is not feasible. Hence, 625 random data samples are selected from the dataset, and visualization is carried out with the Seaborn library. Waist acceleration Y and Waist acceleration Z are the features considered for visualization. In the figure 5, class 0 is ADL, and class 1 is Fall data. Figure 6.2(a) shows the class as the x-axis and the waist acceleration Y as the y-axis. Figure 6.2(b) shows the class as the x-axis and the waist acceleration Z as the y-axis. The visualization indicates waist acceleration Y values for ADL falls between 0 and -20. The waist acceleration Y values for FALL are negative values below -20. In Figure 5.4(b), waist acceleration Z values for ADL are in the range of 0 to -10, and for FALL, the values are more in the positive range from 0 to 10. The visualization clearly shows



(a) Class vs Waist Acceleration Y (b) Class vs Waist Acceleration Z

Fig. 6.2: Data Visualization

the demarcation of the feature values for the two different classes.

7. Conclusions. The proposed accident attention model is developed with the objective of alerting the caregiver to somnambulism when the patient starts sleep walking or has a fall. This accident attention health care system allows the caregiver to sleep without being attentive of the patient or staying awake during the high-risk hours. This is a novel methodology developed by authors to support the caregivers of sleepwalking patients. The study implements only part of the accident attention system, the fall detection module. The future work includes developing the classification model for sleep or wake detection and integrating with the proposed model. The mobile application design for the smart health care of somnambulism patient is also future direction of work.

Funding Statement. This study is supported via funding from Prince Sattam bin Abdul University project number PSAU/2023/R/1444.

Dataset/Data Availability Statement. Robinovitch S, Aziz O, Inertial Measurement Unit Fall Detection Dataset, Federated Research Data Repository, 2017. https://doi.org/10.25314/35954791-2754-47d7-893a-39430535dfdd

Acknowledgments. We would like to acknowledge the assistance of Armaan Ziyad Valiya Parambath in helping conceptualization and editing.

REFERENCES

[1] D. Markov, F. Jaffe, and K. Doghramji, Update on Parasomnias, Psychiatry (Edgmont) 3 (7), 69–76, 2006.
[2] Sleepwalking, nhs.uk, Oct. 18, 2017. https://www.nhs.uk/conditions/sleepwalking/ (accessed Apr. 16, 2023).
[3] Y. Nam, Y. Kim, and J. Lee, Sleep Monitoring Based on a Tri-Axial Accelerometer and a Pressure Sensor, Sensors (Basel), vol. 16, no. 5, p. 750, May 2016, doi: 10.3390/s16050750.
[4] Sleepwalking: What is Somnambulism, Sleep Foundation, Nov. 18, 2020. https://www.sleepfoundation.org/parasomnias/sleepwalking (accessed Apr. 16, 2023).
[5] A. Kales et al., Somnambulism. Clinical characteristics and personality patterns, Arch Gen Psychiatry, vol. 37, no. 12, pp. 1406–1410, Dec. 1980, doi: 10.1001/archpsyc.1980.01780250092011.
[6] Get Enough Sleep - MyHealthfinder | health.gov. https://health.gov/myhealthfinder/healthyliving /mental-health-and-relationships/get-enough-sleep (accessed Apr. 24, 2023).
[7] Future of Smart Health Systems. https://www.news-medical.net/health/Future-of-Smart-Health-Systems.aspx (accessed Apr. 24, 2023).
[8] S. Tian, W. Yang, J. M. L. Grange, P. Wang, W. Huang, and Z. Ye, Smart healthcare: making medical care more intelligent, Global Health Journal, vol. 3, no. 3, pp. 62–65, Sep. 2019, doi: 10.1016/j.glohj.2019.07.001.
[9] S. Shao, G. Han, T. Wang, C. Song, C. Yao, and J. Hou, Obstructive Sleep Apnea Detection Scheme Based on Manually Generated Features and Parallel Heterogeneous Deep Learning Model Under IoMT, IEEE J Biomed Health Inform, vol. 26, no. 12, pp. 5841–5850, Dec. 2022, doi: 10.1109/JBHI.2022.3166859.

[10] L. Haoyu, L. Jianxing, N. Arunkumar, A. F. Hussein, and M. M. Jaber, An IoMT cloud-based real time sleep apnea detection scheme by using the SpO2 estimation supported by heart rate variability, Future Generation Computer Systems, vol. 98, pp. 69–77, Sep. 2019, doi: 10.1016/j.future.2018.12.001.
[11] A. R. Dhruba, K. N. Alam, M. S. Khan, S. Bourouis, and M. M. Khan, Development of an IoT-Based Sleep Apnea Monitoring System for Healthcare Applications, Comput Math Methods Med, vol. 2021, p. 7152576, Nov. 2021, doi: 10.1155/2021/7152576.
[12] K. Saleem, I. S. Bajwa, N. Sarwar, W. Anwar, and A. Ashraf, IoT Healthcare: Design of Smart and Cost-Effective Sleep Quality Monitoring System, Journal of Sensors, vol. 2020, p. e8882378, Oct. 2020, doi: 10.1155/2020/8882378.
[13] F. Yang et al., Internet-of-Things-Enabled Data Fusion Method for Sleep Healthcare Applications, IEEE Internet of Things Journal, vol. 8, no. 21, pp. 15892–15905, Nov. 2021, doi: 10.1109/JIOT.2021.3067905.
[14] Z. Ali, M. S. Hossain, G. Muhammad, and A. K. Sangaiah, An intelligent healthcare system for detection and classification to discriminate vocal fold disorders, Future Generation Computer Systems, vol. 85, pp. 19–28, Aug. 2018, doi: 10.1016/j.future.2018.02.021.
[15] S. Singhal and P. C. Jain, Wireless health monitoring system for sleepwalking patients, in 2015 39th National Systems Conference (NSC), Dec. 2015, pp. 1–6. doi: 10.1109/NATSYS.2015.7489083.
[16] A. Ghubaish, T. Salman, D. Ünal, A. Al-Ali, and R. Jain, Recent Advances in the Internet of Medical Things (IoMT) Systems Security, IEEE Internet of Things Journal, vol. PP, Dec. 2020, doi: 10.1109/JIOT.2020.3045653.
[17] S. R. Ziyad and A. Ziyad, Emergency Rapid Response to Epileptic Seizures-A Novel IOT Framework for Smart Cities, Scalable Computing: Practice and Experience, vol. 22, no. 2, pp. 259–272, 2021.
[18] S. C. Mukhopadhyay, N. K. Suryadevara, and A. Nag, Wearable Sensors for Healthcare: Fabrication to Application, Sensors, vol. 22, no. 14, Art. no. 14, Jan. 2022, doi: 10.3390/s22145137.
[19] J. Srivastava, S. Routray, S. Ahmad, and M. M. Waris, Internet of Medical Things (IoMT)-Based Smart Healthcare System: Trends and Progress, Computational Intelligence and Neuroscience, vol. 2022, p. e7218113, Jul. 2022, doi: 10.1155/2022/7218113.
[20] A. E. Khaled, Internet of Medical Things (IoMT): Overview, Taxonomies, and Classifications, Journal of Computer and Communications, vol. 10, no. 8, Art. no. 8, Aug. 2022, doi: 10.4236/jcc.2022.108005.
[21] M. A. Khelili, S. Slatnia, O. Kazar, and S. Harous, IoMT-fog-cloud based architecture for Covid-19 detection, Biomedical Signal Processing and Control, vol. 76, p. 103715, Jul. 2022, doi: 10.1016/j.bspc.2022.103715.
[22] S. Razdan and S. Sharma, Internet of Medical Things (IoMT): Overview, Emerging Technologies, and Case Studies, IETE Technical Review, vol. 39, no. 4, pp. 775–788, Jul. 2022, doi: 10.1080/02564602.2021.1927863.
[23] Useful Dementia Apps for Seniors and Their Caregivers https://www.aplaceformom.com/caregiver-resources/articles/dementia (accessed Apr. 24, 2023).
[24] The 10 best epilepsy apps. https://www.medicalnewstoday.com/articles/319430 (accessed Apr. 24, 2023).
[25] A. S. Zergham and Z. Chauhan, Somnambulism, in StatPearls, Treasure Island (FL):StatPearls Publishing, 2023. Accessed: Apr. 21, 2023.
[26] O. Walch, Y. Huang, D. Forger, and C. Goldstein, Sleep stage prediction with raw acceleration and photoplethysmography heart rate data derived from a consumer wearable device, Sleep, vol. 42, no. 12, p. zsz180, Dec. 2019, doi: 10.1093/sleep/zsz180.
[27] V. Radha, T. Vayyapuri, and S. Puthu Vedu, Classification of benign and malignant pulmonary nodules in ldct images using radiomic features, Journal of Engineering Science and Technology, vol. 16, pp. 3250–3266, Aug. 2021.
[28] M. Alharbi, Novel Computer-Aided Diagnosis System for the Early Detection of Alzheimer's Disease, Computers, Materials and Continua, vol. 74, pp. 5483–5505, Jan. 2023, doi: 10.32604/cmc.2023.032341.
[29] P. Ghosh et al., Efficient Prediction of Cardiovascular Disease Using Machine Learning Algorithms With Relief and LASSO Feature Selection Techniques, IEEE Access, vol. 9, pp. 19304–19326, 2021, doi: 10.1109/ACCESS.2021.3053759.
[30] Y.Freundand and R. E.Schapire, A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting, Journal of computer and systemsciences, vol. 55,pp. 119139, 1997.
[31] P. Schober and T. R. Vetter, Logistic Regression in Medical Research, Anesthesia and Analgesia, vol. 132, no. 2, p. 365, Feb. 2021, doi: 10.1213/ANE.0000000000005247.
[32] E. Y. Boateng and D. A. Abaye, A Review of the Logistic Regression Model with Emphasis on Medical Research, Journal of Data Analysis and Information Processing, vol. 7, no. 4, Art. no. 4, Sep. 2019, doi: 10.4236/jdaip.2019.74012.
[33] A. Wood, V. Shpilrain, K. Najarian, and D. Kahrobaei, Private naive bayes classification of personal biomedical data: Application in cancer data analysis, Computers in Biology and Medicine, vol. 105, pp. 144–150, Feb. 2019, doi: 10.1016/j.compbiomed.2018.11.018.
[34] K. Taunk, S. De, S. Verma, and A. Swetapadma, A Brief Review of Nearest Neighbor Algorithm for Learning and Classification, in 2019 International Conference on Intelligent Computing and Control Systems(ICCS), May 2019, pp. 1255–1260.
[35] M. Hossin and S. M.N, A Review on Evaluation Metrics for Data Classification Evaluations, International Journal of Data Mining and Knowledge Management Process, vol. 5, pp. 01–11, Mar. 2015, doi: 10.5121/ijdkp.2015.5201.

Edited by: Pradeep Kumar Singh
Received: Apr 25, 2023
Accepted: Jun 2, 2023

Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2023 SCPE. Volume 24, Issues 2, pp. 161–171, DOI 10.12694/scpe.v24i2.2081

A BONE FRACTURE DETECTION USING AI-BASED TECHNIQUES

RUSHABH MEHTA*, PREKSHA PAREEK†, RUCHI JAYASWAL‡, SHRUTI PATIL§ AND KISHAN VYAS¶

Abstract. The medical field in itself is a complex term where the diagnosis is of the most importance. If there is a correct diagnosis made on time in the appropriate time duration then the treatment can be started in a timely manner and this treatment will be beneficial in curing the patient. There are many different techniques that are available to find the abnormalities in an image given but we will review some of them which are most recently developed and will compare the results of each of them. A detailed study is done at the end of this paper which gives insights into fractures and their types. The dataset which we would consider is the MURA dataset. Discussion about further research in this area is also done to help researchers in exploring new dimensions in this field.

Key words: Machine Learning, Artificial intelligence, Bone Fractures, Medical Images, X-Rays, CAD.

1. Introduction. The adult human body consists of 206 bones, these include different types of bones which are different in shape, size, structure and morphology. There are many causes of fracture in human bone these may be due to old age, accident, or falling. The risk of getting a bone fracture mostly increases in such cases when the bones are weak and there are deficiencies of required vitamins, minerals and calcium [32]. The most used method for the interpretation of fracture in humans is an x-ray or medical resonance imaging [37]. Sometimes there are cases when the cracks in a bone are very small and doctors cannot detect them easily so there is a need for enhancement in the diagnosis of bone fracture detection with the help of computer-aided diagnosis or machine learning coupled with different methods [4].

The normally or usually used method which is used to find fracture in a bone is by viewing it by a radiologist who must use his experience in inspecting the X-ray image and giving the results after visually inspecting it. Most of the time technicians have to take x-ray images in different poses like anterior view, posterior view and lateral view as many times it happens that there are cases where by viewing different views like anterior view or posterior view radiologist cannot come to one given results in such cases, he has to use different views of x-rays so that he can give a result if a bone is a fracture or not. But this is not sufficient in diagnosis as sometimes a doctor may miss a diagnosis and, in such cases, there may be chances that the patient may have to undergo a prolonged treatment.

In some cases, there are many times patient ignores some symptoms and the diagnosis by a doctor is done based on their complaints as many times the patient gives incomplete information about his/her health condition. Sometimes there are cases where there may be chances that the x-ray image is not taken properly in such cases also there are chances of missing the diagnosis. These are the cases of bone fracture. But in complex cases where there is a need for a CT scan the whole diagnosis is mostly based on the complaints. Thus, in this survey, the main motive is to help the medical fraternity in diagnosing x-ray images.

This paper could further help- researchers to improve the given technology and will motivate them in finding

*Department of Artificial Intelligence and Machine Learning, Symbiosis Institute of Technology, Pune, India (rushabh.mehta.mtech2021@sitpune.edu.in)
†Department of Artificial Intelligence and Machine Learning, Symbiosis Institute of Technology, Pune, India (preksha.pareek@sitpune.edu.in)
‡Department of Artificial Intelligence and Machine Learning, Symbiosis Institute of Technology, Pune, India (ruchi.jayaswal@sitpune.edu.in)
§Department of Artificial Intelligence and Machine Learning, Symbiosis Institute of Technology, Pune, India (shruti.patil@sitpune.edu.in)
¶Department of Artificial Intelligence and Machine Learning, Symbiosis Institute of Technology, Pune, India (kishan.vyas.mtech2021@sitpune.edu.in)

Fig. 1.1: Sample X-Ray Image

new ways in forming more accurate systems. X-ray images which are considered by different authors are from different sources and the accuracy is mentioned in the cases where results were available. Further, there is a model which is proposed which could give more accurate results. These technologies of machine learning and computer-aided diagnosis can be further used in more complex diagnoses like CT scans or MRIs.

The process of first diagnosing the fracture is time-consuming with there being the difficulty of radiologist experts present in villages or remote areas. Thus, with the help of machine learning there can be adverse changes that can be made for effective diagnosis of x-rays [10, 12]. In the past decade, there has been an increasing application of convolution neural network models for the detection of bone fractures, these methods have demonstrated that the use of deep learning has also proved to be beneficial and has added advantage for this purpose [27, 40]. Figure 1.1 represent sample of X-Ray image.

2. Motivation. X-ray with auto-report generation by the machine itself, the day is not far when there would be ultrasound and these reports would be analyzed by the machine itself based on the data by which it is trained. There can be a machine learning model by which the machines are trained and they could be trained to such an extent that they could predict far better than doctors. The radiologist would visit only twice or thrice a week, and ultrasound, X-ray reporting and MRI could be performed on these days. For any critical illness, this type of infrastructure would be dangerous and if these types of machines come up then a technician could perform X-rays and reports would be analyzed by the machine itself (maybe the patient could act faster). This could not only be beneficial to rural areas but also could be faster and more precise in the detection of diseases.

Bone is an important part of the human body, without bones, we would not be able to walk, run, jump, climb stairs, or do any physical activity. bones also help us to carry out our daily activities. when a person falls down, he/she can break his/her bones. to make a diagnosis, they must know how to read these images. they then compare them to previous ones to see if the same injury has happened before. for example, if a patient had a broken arm last year, the doctor might ask about the location of the injury. he/she might also ask about the type of injury (fracture vs. dislocation).

Finally, the doctor will check whether the injury was caused by a fall or by something else. hence our main goal is to make a robust system that can help medical professionals in interpreting x-ray images and help them identify bone fractures. There are mainly two methods that we can use to primarily develop a model these include machine learning and deep learning.

2.1. Contribution of the work. A contribution of this review paper as follows:

- X-ray images of the bone are taken with medical instruments, x-ray is mostly used for the detection of bone fractures. Bone fractures are those abnormalities that are caused when a given bone cannot withstand the given pressure from outside and due to this, there is the development of cracks in bones.

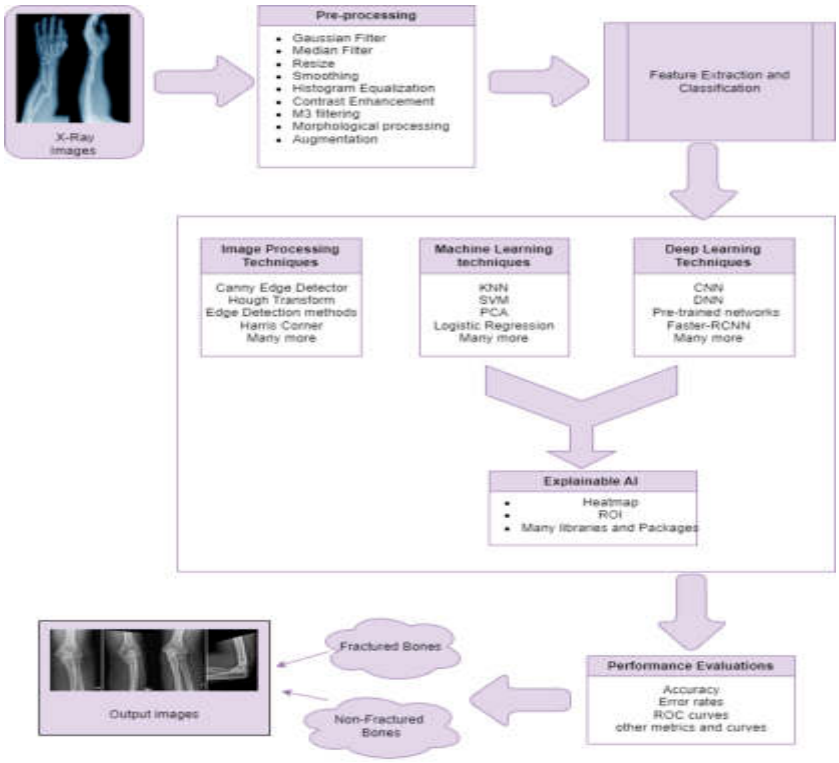


Fig. 3.1: Block Diagram of Bone Fracture Detection

- It is of utmost importance for doctors to detect bone fractures in time to ensure timely treatment is provided to the patient.
- Thus, new technologies and advancements in artificial intelligence and machine learning there have new reasonable ways to detect bone fractures on different types of bones.

2.2. *Organization of the paper*. Section 2 defines the related work carried out previously by different researchers. Section 3 defines the different datasets which are openly available for this research. Section 4 defines the evaluation metrics. The further part section 5 details the challenges which have to be addressed. Section 6 gives a conclusion about our learnings from this paper. Section 7 narrates the further scope of this research and the further enhancement of this field.

3. Background Work. In the proposed model we acquire the images from the MURA dataset and select one category of x-ray from the seven types of bone X-rays such as the forearm, elbow, wrist, humerus, hand, fingers, and shoulder. The next step involves the pre-processing of these images, these methods include some of the mentioned as a gaussian filter, median filter, smoothing, resizing, contrast enhancement, histogram equalization and augmentation. After the pre-processing is done, we move towards feature extraction and classification. There are machine learning techniques as well as deep learning techniques which are involved, we may use any one of them which fetches better performance. Machine learning technique involves the use of SVM, KNN, Logistic Regression, etc. While deep learning techniques involve CNN, Faster R-CNN, DNN, etc. with the use of these techniques our model will classify x-ray image as fractured and normal images with a certain accuracy level. We now will use techniques like ROI, and heatmap to find and explain the behaviours of our model. We move towards performance evaluation of our model, these include accuracy, precision, recall, AUC, ROC, etc. after the performance is evaluated, we try to change parameters and fine tune it in order to get better performance.

The study from [4] uses primitive machine learning techniques such as feature extraction and pre-processing.

Developed the RDSS method to recover incorrect parameters to overcome errors in contour segmentation [5]. The same author has developed in [6] a fracture diagnosis model based on machine learning techniques. Initially, the gaussian filter is used to enhance the quality of the given X-ray image. The corners and edges are then identified using the Canny Edge algorithm. Lastly, the Harris corner detection method is used to detect the fractured areas. The method can classify fractured bones with 92% accuracy. Feature extraction is enhanced by various image processing techniques, such as pixel density, controllable filters, and image projection integration [17].

All of the above-mentioned methods do an excellent job of diagnosing the skeletal system. However, the performance of these techniques is not satisfactory due to the handmade features used to train the model. Recently, some authors reported that adding deep convolution neural network models to a set of, features gave improved classification performance. More recently, the authors of [26] developed the Regional Convolutional Neural Network (R-CNN) for the detection of skull fractures. They leveraged the prior clinical knowledge of the fastest R-CNN to improve the classification performance.

The authors of the paper [22] have developed a deep-ensembled-based convolution neural network(CNN) model for ankle fracture detection. Their methods include ResNet and InceptionV3 for feature extraction. The cluster-based model classified healthy and fractured individuals with 81% accuracy. Among transfer-based learning methods, the authors of [11] have also used the YOLOV3 model was used to identify skull fractures which had a sensitivity of 91.7%. In, subsequent research [21], pre-trained InceptionV3 and DenseNet-121 models were used for bone diagnosis. The methods mentioned in the paper [21] achieved an accuracy of 95%, while the method described in the paper [20] achieved an accuracy of 95.4%. Some studies reported that they first trained the models on a dataset of skeletal images before classification.

The authors of [45] have developed two deep convolutions neural network(CNN) models for the identification and segmentation of intertrochanteric fractures. The dataset is divided into two parts, training, and testing, with 32,045 and 11,465 images, respectively. First, a region of interest (ROI) is identified using a CNN based on a cascade structure. Another CNN is then used for segmentation and recognition. In other studies, of paper considered the researchers of the paper [45] have pre-processed rib fracture images, and resized the image from 128 x 128 x 333 pixels. The semantic segmentation technique was later used to locate the fractured regions of the ribs. Finally, a UNet model is used to classify the CT images with an accuracy of 88.5%. in the paper [33] the authors have used YOLO Model for bone localization of fractures. Data augmentation techniques are also applied and the performance of their models is compared with the original datasets. In the original dataset, the method can classify fractured and non-fractured regions with an accuracy of 81.9

In the paper [42] researchers have designed a hierarchical network and compared it to orthopedic diagnoses. They trained hierarchical matrices on X-ray images. Their method achieved a classification accuracy of 88.7%. In the research paper [30] use fractures were identified in two stages; In the initial step, the fastest R-CNN was used to locate 20 fracture regions, and in the next step, the new CrackNet network was used to classify the fractures. Their method classified healthy bones and, fractured bones with a classification accuracy of 90.1%. In a similar study, a new method was proposed to classify the parallel net fraction using the two-scale method proposed by [41].

The authors of [44, 3] used pre-trained R-CNN was applied to the small dataset and achieved 96% and 97% accuracy, respectively. The authors of [29] have applied their experiences and designed a decision tree for the identification of fractures. Their method achieved a classification accuracy of, of 86.57%. The researchers of [8, 19] have used deep CNN applications to extract features from skeletal images. The validation accuracy of their method is 83% and 97.4%, respectively. The authors of [14] and Rayvolve model was designed to detect, fractures in children. Externally validated methods classified healthy and broken bones with 95% accuracy.

In 2022 the authors of [16] used a deep CNN model based on clusters was applied to determine wrist fractures. Their method isolated, broken, and healthy wrist bones with 86.39% accuracy. The authors of [34] have used in 2019 classified bone fractures by combining, depth lines and SURF features. They compared the performance of ResNet and VGG16 among which, deep CNN models have been previously trained. Of the two models considered the ResNet model has the highest classification accuracy of 98%.

In 2020 the researchers of [31] have used techniques like Generative Adversarial Networks and Digitally Reconstructed Radiographs. The image dataset was generated using GAN and a deep CNN model was used to

Table 3.1: Literature Review of Bone Fracture Detection by Computer Vision Techniques

Year	Methods	Pros	Cons	Datasets	Performance
2021 [35]	KNN classification	Accuracy is good to rely on for the detection of osteoporosis.	As the feature dimension increases. the accuracy suffers. Has been implemented for the lumbar spine only.	Self-synthesized dataset.	97.22% Accuracy
2019 [36]	Image Pre-Processing: 1)Smoothing 2)Histogram equalizer 3)Edge detection 4)Segmentation 5)Contrast enhancement	The theory of enhancement of images and then identifying the focus area using computer vision is used In this paper.	The dataset is not well defined. Accuracy is not declared in this paper.	The dataset is collected from different medical institutes and one dataset is formed from the obtained data.	Not Available
2020 [37]	Harris Corner method is used for the detection of fracture and for image pre-processing M3 filtering is used as it gives good PSNR(peak signal noise ratio) compared to mean filtering and . median filtering.	It gives us new direction for us in this field as these new methods can be used for fracture detection rather than using CNN or Deep learning.	The dataset is self-synthesized so it is difficult to procure this dataset.	Self-synthesized dataset.	94% Accuracy
2018	Morphological operations, Special Bone Feature, Extraction, Sobel Edge Detection	Hand fingers are only used thus which helps in increasing the accuracy.	Hand fingers are only used but it does not specify properly which finger is used like the thumb,index, etc.	Self-synthesized dataset(155 images (100 fractures images, 30 hands finger images, 25 normal images)	92% for general fracture, 93.33% for finger
2018 [38]	Gaussian filter, canny edge detector, Sobel Threshold	It uses the technique of the Gaussian filter which has given accurate results.	The dataset is self-synthesized so it is difficult to procure this dataset.	Self-synthesized dataset.	Not Defined
2017 [39]	Canny edge detector, Hough Transform	New methods of the Hough transform are used for bone fracture detection.	A small dataset of only 10 images is used due to which . there is a lack of training in the model.	10images	80% Accuracy
2017 [40]	Gaussian Filter, Gradient Magnitude, Canny Detector	Methods of computer vision give new direction to this work.	A very small dataset of only 12 images is used for training the model.	12images	Not Defined
2016 [41]	Canny detector.	It only focuses on one technique to detect fractured images thus gaining a good amount of accuracy.	The dataset used is an imbalanced dataset thus it should first balance the dataset.	Self-synthesized dataset. 21 images (16 images are normal 5 images are fractured)	87.5% Accuracy
2016 [42]	Morphological Gradient, Smoothing, Canny Detector	Methods of gradient and canny detector are used which gives new direction to research.	The dataset used is not mentioned. The accuracy or any other metric to evaluate is not defined.	Not defined.	Not defined.
2015 [43]	Laplacian Gradient, K means clustering	Techniques used are based on computer vision thus making new methods available for the purpose of bone fracture detection.	The dataset used is not defined.	Not defined.	85%
2013 [44]	Edge detection, texture detection, parallel edges	Edge detection along with parallel edges is used.	The dataset used is an imbalanced dataset thus it needs to be first balanced and then used so that training would give accurate results.	Self-synthesized dataset 300 (200 normal, 100 fractures images)	83%

perform the classification. The researchers of [43] have applied a CNN model based on deep learning, obtaining an accuracy of 90.2% to identify rib fractures. The authors of [38] have used the InceptionV3 model to classify radiographs of the proximal femur. Their method classified, images of healthy and broken bones with 86 percent accuracy. In a similar study, the researchers of [28] have uses DenseNet to classify femoral fractures with 89% accuracy.

3.1. Bone Fracture Detection using Computer Vision. Generally used methods are based on concepts of computer vision and image processing. Computer vision has an immense application in the area of the medical field and has been proven helpful with reasonable performance levels. Some of these methods also use concepts of computer vision along with machine learning or deep learning. The use of computer vision in x-ray bone fracture detection is used for more than a decade and researchers are continuously enhancing and introducing new methods for better performances.

Table 3.1 compares some different other methods and their application for the purpose of bone fracture detection. These papers included are from 2012 to the present.

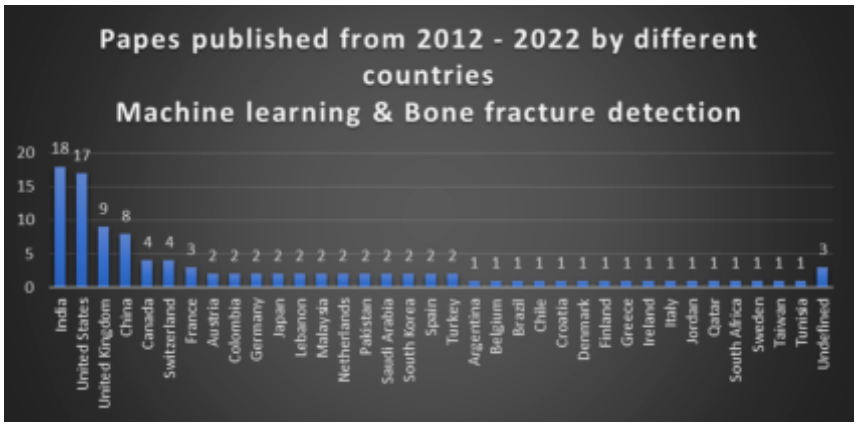


Fig. 3.2: Country-wise papers for ML and bone fracture detection

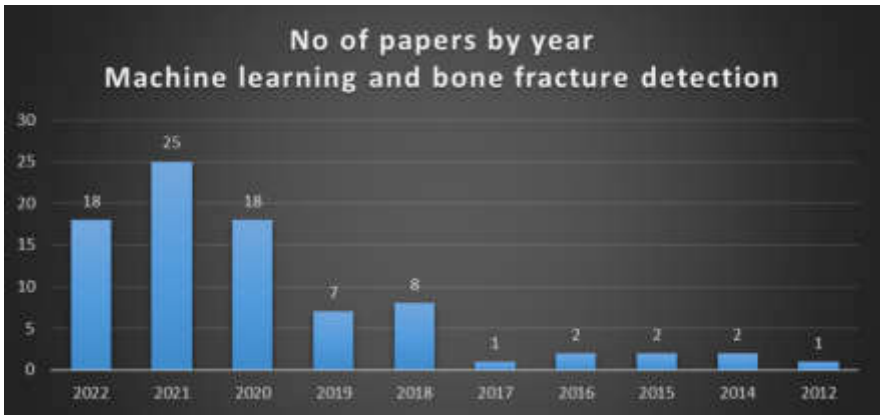


Fig. 3.3: Year-wise paper for ML and bone fracture detection

3.2. Bone Fracture Detection using Machine Learning. Bone fracture using machine learning in itself has a very wide scope due variety of methodologies available for this purpose. There is a variety of different classification algorithms which can be used in order to classify healthy bone and fractured bone. Different classification methods involve KNN [7], SVM [7], Logistic Regression [23], Naïve Bayes, etc are few to mention. Machine learning methods have proven to give acceptable accuracy levels for purpose of bone fracture detection. From the graphs and the data collected from different data sources, it can be seen that the use of machine learning applied for the purpose of bone fracture detection has seen a declining trend after 2020 and 2021 as the concept and use of deep learning methods have been proving more advantages in both simple computations, less computation time as well as a higher level of accuracy.

Figure 3.2 depicts and gives insight into the number of papers published for bone fracture detection using machine learning from 2012 to 2022 by different countries. The data are collected from Scopus.

Figure 3.3 depicts and gives insight into the number of papers published for bone fracture detection using machine learning in the last decade. The data are collected from Scopus.

3.3. Bone Fracture Detection using Deep Learning. Bone fracture detection using deep learning involves different techniques such as Convolution Neural Networks [36], Artificial Neural Networks, Deep Neural Networks [36], Deep Convolution Neural Networks [27], Inception V3 [39], R-CNN, etc few to mention. The concept of use and application of deep learning towards the use of bone fracture detection has taken momentum

after 2019 as these methods or techniques once tried have given very interesting and useful results. The drawbacks of previously used different methods have been overcome using the concept of deep learning. There was a problem with analysing images from only one anatomical position and due to this, there are some cases where there is missed diagnosis. Due to this deep learning method and the use of faster R-CNN and customized cracknet [30] the minute cracks in the x-rays were been able to be detected.

3.4. Bone fracture detection using Explainable AI. Explainable AI(XAI) [46] refers to a technique in which we try to explain why a machine learning algorithm or deep learning algorithm has reached a solution. In our case, we consider some papers which try to explain why the model has predicted the outcome as normal bone or fractured bone. The XAI has been in use for the explainability of the model and the research on why the model has reached some definite conclusion.

In the paper [18] authors have used classification techniques for finding fractured and normal bones. They have used full radiographs as well as the manually defined region of interest to help classify better. In these cases, the accuracy of full radiographs was 835 whereas the accuracy level of manually decided regions of interest the accuracy level was 93%. The accuracy level was the same for automatic localized x-ray images 93%. In the paper [46] researchers used binary classification on the whole dataset and managed to get 96.9% accuracy, use of the CNN algorithm fetched different results for different bone types. The authors of [9] have used deep learning systems for the purpose of bone fracture classification and have achieved an AUC OF 0.93 and a 95% confidence interval. It was found that in around 90% of x-ray images the model prediction was in line with radiologist annotation. The authors of [24] have used densely connected convolution neural networks and have achieved an accuracy of 95 .8%, they have further used binary classification for the femoral neck (displaced and non-displaced), and intertrochanteric fracture. In the paper [27] authors have used deep learning algorithms with aided and unaided diagnosis and have found an AUC of 0.967. Explainable AI tries to answer some of the most important questions such as what are the model's weaknesses and model strengths, the criteria by which the model has chosen to arrive at a specific conclusion, the most important why has model chosen this approach and given particular solution as opposed to other, what are the errors that may be faced by certain models and how to correct these errors is all that explainable AI explains.

4. Datasets. The most openly available and free-access datasets for this research are as follows:

MURA Dataset. MURA [35] dataset is a large collection of x-ray images from a huge number of patients collected over a large period of time. The samples which are collected include one of the seven types of bones such as forearm, elbow, wrist, humerus, hand, fingers, and shoulder. This huge dataset is collected from 12,173 patients which contains a total of 14,863 x-ray images of bone with a total multi-view of the different anatomical positions of 40,561 x-ray images.

Medpix Dataset. Medpix [25] is an open-source database that is available to the public. It has a vast variety of data including cases in every vertical. The database collected is from nearly about 12,000 patients and it has around 59,000 images in 9000 topics.

Imaging Archive (TCIA). TCIA [13] is a dataset of a collection of cancer images. The cancer imaging archive has a very large and free source of the dataset which is available to the public. It is mainly funded by CIP a part of the National cancer institute which is managed by FNLCR. Figure 10 is a sample image from TCIA which indicates both the chest x-ray as well as CT image of the chest of the same patient.

Radiopedia Dataset. Radiopedia [2] started in 2007 has been the most trusted and reliable source of datasets for radiology which is available for free to the public. In 2021 radiopedia has achieved to serve and provide data to almost 41 million people. Every single country on earth has in some or another other way been provided data from radiopedia. Figure 4.1 shows the frontal and lateral view of the right hand having an intra-articular radial styloid fracture.

Alyward Dataset. Alyward.org [1] has a dataset of chest x-ray images. There is a vast collection of 10,000 chest x-ray images along with the diagnosis.

Diagnostic Imaging Dataset (DID). The dataset [15] is a central collection of imaging that is carried out every month on NHS patients, and it is then extracted locally from the radiology information center. The diagnostic imaging dataset is only available through the data access request service. Other datasets which are available are Dataset by the Institute of Engineering Science and Technology, Shibpur (IIEST).

Fig. 4.1: Sample Image from Radiopedia Dataset

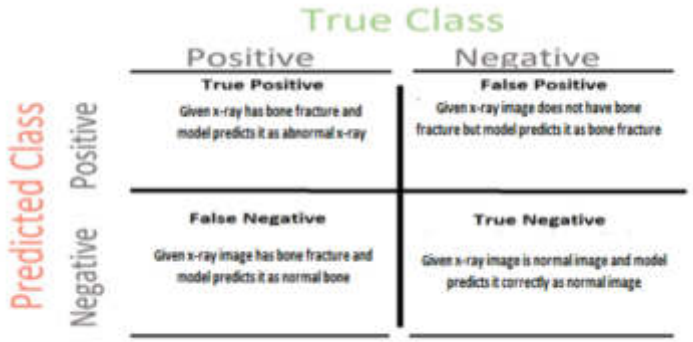


Fig. 5.1: Confusion matrix for bone fracture detection

5. Evaluation Metrics. The evaluation measures of the models are calculated using accuracy, recall, precision and F1 score. These metrics are evaluated based on the confusion matrix of each model using the mathematical formula.

Confusion matrix it is used to give details of the total performance of the classifier model. Accuracy can sometimes portray wrong or false results in cases where the dataset is considered to be imbalanced or uneven dataset. An uneven dataset is a dataset where there are an unequal number of samples for each category. Sometimes classification accuracy can give false results when there are more than two categories being considered. By evaluating the confusion matrix researchers get to understand better what is true and what is false in the model. In the confusion matrix, the correct and incorrect cases are grouped together and their total number of counts is divided by a particular category. The confusion matrix basically tries to portray how confused the classifier is when classifying the datasets. It tries to identify the errors which the classifier makes and also helps in identifying types of errors made by the classifier. Figure 5.1 defines the different aspects of true positive, true negative, false positivea and false negative in respect of bone fracture detection.

Accuracy can be defined as the ratio of total correct predictions to the total no of samples considered.

$$\text{Accuracy} = (TP+TN)/\ (TP+TN+FP+FN) \quad (1)$$

Precision can be defined as the ratio of actual results obtained to that of positive samples.

$$\text{Precision} = TP/(TP+FP) \quad (2)$$

Recall can be defined as the ratio of the total number of positive cases to the total number of predictions done.

Recall = TP/ (TP+FN) (3)

F1-Score measures the harmonic mean of the model performance.

F1 − Score = (2 × (Precision × Recall)) / (Precision + Recall) (4)

6. Challenges. There have been various challenges which are faced while working on research related to the medical field, as there has been a huge number of problems related to the acceptance of new systems and public acceptance is of most importance as they decide the fate of the success of the research. Some of the challenges are mentioned below:

Generalized model: In general, an adult human has 206 bones. Every bone in the human body has a different shape, size, structure and different morphological features. To generalize and form a model which is a general model which can classify between fractured and non-fractured images for any given bone is practically not possible as in such cases for every type of bone there must be training and the accuracy level would also differ for every type of bone.

To form a model which can generalize every bone would take a huge amount of dataset and the challenge would be to gather a huge amount of dataset as for specific types of bone finding both fractured and normal images is difficult and these datasets are most of the time, not open source.

Time complexity and large dataset: If one gets all images by collaborating with an imaging center or hospital the training of such models will also take a huge amount of time and computation complexity would also be very large.

Integration with machines: Integrating these models with x-ray machines for spot diagnosis is also a key challenge as it requires a change of the whole system.

Acceptance by the general public is a key challenge as in cities there are adequate amounts of resources and doctors available for diagnosis but in villages, there is a lack of resources and doctors thus there these types of models can be proved to be a boon to mankind for its spot diagnosis as well as lower cost of diagnosis.

7. Conclusion. This paper contains a full end-to-end analysis of the considered papers and the techniques used in this paper. Methodologies used in these papers and detailed pros and cons have been discussed. The paper also discusses and compares the accuracy level of the considered papers. In the results and discussion, the new age problems related to these technologies and methodologies have been discussed in detail. Thus, the advancement in the field of medical science due to the extensive use of machine learning can be seen in the above learnings. It can be often learned that the use of machine learning towards such a humanitarian cause can be very useful in developing a stronger healthcare infrastructure in the country. It has been seen that the use of machine learning has become obsolete after the introduction of the use of deep learning for the purpose of bone fracture detection as the level of accuracy being achieved by deep learning is commendable and has seen practical life applications. The use of such technology and integration with the medical field should be encouraged at a national level for the technology to reach remote areas as well as encourage new researchers to enhance these types of projects. Further direction to this work is that researchers have used different techniques to develop models and to classify if a given x-ray image is a fractured or normal image. This work can be further given more impact by classifying the types of fractures i.e., stress fracture, oblique fracture, open fracture, close fracture, compression fracture, etc. In this further work, researchers can try to combine different methodologies from the referred papers and try to form a model which can give the highest level of accuracy and can be used by medical professionals in their routine. The other aspect that researchers can try to achieve is to integrate their formed model with X-Ray machines.

REFERENCES

[1] *Alywarddataset*. `https://nihcc.app.box.com/v/ChestXray-NIHCC/folder/37178474737`.

[2] *Radiopediadataset*. `https://radiopaedia.org/search?lang=us&q=fracture&scope=cases`.

[3] W. Abbas, S. M. Adnan, M. A. Javid, W. Ahmad, and F. Ali, *Analysis of tibia-fibula bone fracture using deep learning technique from x-ray images*, International Journal for Multiscale Computational Engineering, 19 (2021).

[4] T. Anu and R. Raman, *Detection of bone fracture using image processing methods*, Int J Comput Appl, 975 (2015), p. 8887.

[5] O. Bandyopadhyay, A. Biswas, and B. B. Bhattacharya, *Long-bone fracture detection in digital x-ray images based on digital-geometric techniques*, Computer methods and programs in biomedicine, 123 (2016), pp. 2–14.

[6] C. Z. Basha, M. R. K. Reddy, K. H. S. Nikhil, P. Venkatesh, and A. Asish, *Enhanced computer aided bone fracture detection employing x-ray images by harris corner technique*, in 2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC), IEEE, 2020, pp. 991–995.

[7] S. M. Beram, H. Pallathadka, I. Patra, and P. Prabhu, *A machine learning based framework for preprocessing and classification of medical images*, ECS Transactions, 107 (2022), p. 7589.

[8] S. Beyaz, K. Açici, and E. Sümer, *Femoral neck fracture detection in x-ray images using deep learning and genetic algorithm approaches*, Joint diseases and related surgery, 31 (2020), p. 175.

[9] C. Blüthgen, A. S. Becker, I. V. de Martini, A. Meier, K. Martini, and T. Frauenfelder, *Detection and localization of distal radius fractures: Deep learning system versus radiologists*, European journal of radiology, 126 (2020), p. 108925.

[10] A. Brett, C. G. Miller, C. W. Hayes, J. Krasnow, T. Ozanian, K. Abrams, J. E. Block, and C. van Kuijk, *Development of a clinical workflow tool to enhance the detection of vertebral fractures: accuracy and precision evaluation*, Spine, 34 (2009), pp. 2437–2443.

[11] J. W. Choi, Y. J. Cho, J. Y. Ha, Y. Y. Lee, S. Y. Koh, J. Y. Seo, Y. H. Choi, J.-E. Cheon, J. H. Phi, I. Kim, et al., *Deep learning-assisted diagnosis of pediatric skull fractures on plain radiographs*, Korean Journal of Radiology, 23 (2022), p. 343.

[12] S. W. Chung, S. S. Han, J. W. Lee, K.-S. Oh, N. R. Kim, J. P. Yoon, J. Y. Kim, S. H. Moon, J. Kwon, H.-J. Lee, et al., *Automated detection and classification of the proximal humerus fracture by using deep learning algorithm*, Acta orthopaedica, 89 (2018), pp. 468–473.

[13] K. Clark, B. Vendt, K. Smith, J. Freymann, J. Kirby, P. Koppel, S. Moore, S. Phillips, D. Maffitt, M. Pringle, et al., *The cancer imaging archive (tcia): maintaining and operating a public information repository*, Journal of digital imaging, 26 (2013), pp. 1045–1057.

[14] M. Dupuis, L. Delbos, R. Veil, and C. Adamsbaum, *External validation of a commercially available deep learning algorithm for fracture detection in children*, Diagnostic and Interventional Imaging, 103 (2022), pp. 151–159.

[15] N. England, *Diagnostic imaging dataset*, 2014.

[16] F. Hardalaç, F. Uysal, O. Peker, M. Çiçeklidağ, T. Tolunay, N. Tokgöz, U. Kutbay, B. Demirciler, and F. Mert, *Fracture detection in wrist x-ray images using deep learning-based object detection models*, Sensors, 22 (2022), p. 1285.

[17] N.-D. Hoang and Q.-L. Nguyen, *A novel method for asphalt pavement crack classification based on image processing and machine learning*, Engineering with Computers, 35 (2019), pp. 487–498.

[18] A. Jiménez-Sánchez, A. Kazi, S. Albarqouni, C. Kirchhoff, P. Biberthaler, N. Navab, S. Kirchhoff, and D. Mateus, *Precise proximal femur fracture classification for interactive training and surgical planning*, International journal of computer assisted radiology and surgery, 15 (2020), pp. 847–857.

[19] R. M. Jones, A. Sharma, R. Hotchkiss, J. W. Sperling, J. Hamburger, C. Ledig, R. O'Toole, M. Gardner, S. Venkatesh, M. M. Roberts, et al., *Assessment of a deep-learning system for fracture detection in musculoskeletal radiographs*, NPJ digital medicine, 3 (2020), p. 144.

[20] D. Kim and T. MacKinnon, *Artificial intelligence in fracture detection: transfer learning from deep convolutional neural networks*, Clinical radiology, 73 (2018), pp. 439–445.

[21] G. Kitamura, *Deep learning evaluation of pelvic radiographs for position, hardware presence, and fracture detection*, European journal of radiology, 130 (2020), p. 109139.

[22] G. Kitamura, C. Y. Chung, and B. E. Moore, *Ankle fracture detection utilizing a convolutional neural network ensemble implemented with a small sample, de novo training, and multiview incorporation*, Journal of digital imaging, 32 (2019), pp. 672–677.

[23] S. H. Kong, D. Ahn, B. Kim, K. Srinivasan, S. Ram, H. Kim, A. R. Hong, J. H. Kim, N. H. Cho, and C. S. Shin, *A novel fracture prediction model using machine learning in a community-based cohort*, JBMR plus, 4 (2020), p. e10337.

[24] J. D. Krogue, K. V. Cheng, K. M. Hwang, P. Toogood, E. G. Meinberg, E. J. Geiger, M. Zaid, K. C. McGill, R. Patel, J. H. Sohn, et al., *Automatic hip fracture identification and functional subclassification with deep learning*, Radiology: Artificial Intelligence, 2 (2020), p. e190023.

[25] N. library of medicine, *Medipix*. `https://medpix.nlm.nih.gov/search?allen=true&allt=true&alli=true&query=fracture`.

[26] X. Lin, Z. Yan, Z. Kuang, H. Zhang, X. Deng, and L. Yu, *Fracture r-cnn: An anchor-efficient anti-interference framework for skull fracture detection in ct images*, Medical Physics, (2022).

[27] R. Lindsey, A. Daluiski, S. Chopra, A. Lachapelle, M. Mozer, S. Sicular, D. Hanel, M. Gardner, A. Gupta, R. Hotchkiss, et al., *Deep neural network improves fracture detection by clinicians*, Proceedings of the National Academy of Sciences, 115 (2018), pp. 11591–11596.

[28] M. Lotfy, R. M. Shubair, N. Navab, and S. Albarqouni, *Investigation of focal loss in deep learning models for femur fractures classification*, in 2019 International Conference on Electrical and Computing Technologies and Applications (ICECTA), IEEE, 2019, pp. 1–4.

[29] J. Luo, G. Kitamura, E. Doganay, D. Arefan, and S. Wu, *Medical knowledge-guided deep curriculum learning for elbow fracture diagnosis from x-ray images*, in Medical Imaging 2021: Computer-Aided Diagnosis, vol. 11597, SPIE, 2021, pp. 247–252.

[30] Y. Ma and Y. Luo, *Bone fracture detection through the two-stage system of crack-sensitive convolutional neural network*, Informatics in Medicine Unlocked, 22 (2021), p. 100452.

[31] S. Mutasa, S. Varada, A. Goel, T. T. Wong, and M. J. Rasiej, *Advanced deep learning techniques applied to automated femoral neck fracture detection and classification*, Journal of Digital Imaging, 33 (2020), pp. 1209–1217.

[32] S. Myint, A. S. Khaing, and H. M. Tun, *Detecting leg bone fracture in x-ray images*, International Journal of Scientific & Research, 5 (2016), pp. 140–144.

[33] H. P. Nguyen, T. P. Hoang, and H. H. Nguyen, *A deep learning based fracture detection in arm bone x-ray images*, in

2021 international conference on multimedia analysis and pattern recognition (MAPR), IEEE, 2021, pp. 1–6.
[34] Y. D. Pranata, K.-C. Wang, J.-C. Wang, I. Idram, J.-Y. Lai, J.-W. Liu, and I.-H. Hsieh, *Deep learning and surf for automated classification and detection of calcaneus fractures in ct images*, Computer methods and programs in biomedicine, 171 (2019), pp. 27–37.
[35] P. Rajpurkar, J. Irvin, A. Bagul, D. Ding, T. Duan, H. Mehta, B. Yang, K. Zhu, D. Laird, R. L. Ball, et al., *Mura: Large dataset for abnormality detection in musculoskeletal radiographs*, arXiv preprint arXiv:1712.06957, (2017).
[36] A. Solovyova and I. Solovyov, *X-ray bone abnormalities detection using mura dataset*, arXiv preprint arXiv:2008.03356, (2020).
[37] B. Swathika, K. Anandhanarayanan, B. Baskaran, and R. Govindaraj, *Radius bone fracture detection using morphological gradient based image segmentation technique*, Int J Comput Sci Inf Technol, 6 (2015), pp. 1616–1619.
[38] L. Tanzi, E. Vezzetti, R. Moreno, A. Aprato, A. Audisio, and A. Massè, *Hierarchical fracture classification of proximal femur x-ray images using a multistage deep learning approach*, European journal of radiology, 133 (2020), p. 109373.
[39] Y. Thian, Y. Li, P. Jagmohan, D. Sia, V. Chan, and R. Tan, *Convolutional neural networks for automated fracture detection and localization on wrist radiographs. radiol artif intell 1: e180001*, 2019.
[40] T. Urakawa, Y. Tanaka, S. Goto, H. Matsuzawa, K. Watanabe, and N. Endo, *Detecting intertrochanteric hip fractures with orthopedist-level accuracy using a deep convolutional neural network*, Skeletal radiology, 48 (2019), pp. 239–244.
[41] M. Wang, J. Yao, G. Zhang, B. Guan, X. Wang, and Y. Zhang, *Parallelnet: Multiple backbone network for detection tasks on thigh bone fracture*, Multimedia Systems, (2021), pp. 1–10.
[42] M. Wang, G. Zhang, B. Guan, M. Xia, and X. Wang, *Multiple reception field network with attention module on bone fracture detection task*, in 2021 40th Chinese Control Conference (CCC), IEEE, 2021, pp. 7998–8003.
[43] T. Weikert, L. A. Noordtzij, J. Bremerich, B. Stieltjes, V. Parmar, J. Cyriac, G. Sommer, and A. W. Sauter, *Assessment of a deep learning algorithm for the detection of rib fractures on whole-body trauma computed tomography*, Korean journal of radiology, 21 (2020), p. 891.
[44] E. Yahalomi, M. Chernofsky, and M. Werman, *Detection of distal radius fractures trained by a small set of x-ray images and faster r-cnn*, in Intelligent Computing: Proceedings of the 2019 Computing Conference, Volume 1, Springer, 2019, pp. 971–981.
[45] L. Yang, S. Gao, P. Li, J. Shi, and F. Zhou, *Recognition and segmentation of individual bone fragments with a deep learning approach in ct scans of complex intertrochanteric fractures: A retrospective study*, Journal of Digital Imaging, (2022), pp. 1–9.
[46] J. Yu, S. Yu, B. Erdal, M. Demirer, V. Gupta, M. Bigelow, A. Salvador, T. Rink, S. Lenobel, L. Prevedello, et al., *Detection and localisation of hip fractures on anteroposterior radiographs with artificial intelligence: proof of concept*, Clinical radiology, 75 (2020), pp. 237–e1.

Edited by: Katarzyna Wasielewska
Received: Dec 10, 2023
Accepted: Jun 11, 2023


Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2023 SCPE. Volume 24, Issues 2, pp. 173–180, DOI 10.12694/scpe.v24i2.2102

EVALUATING THE IGRAPH COMMUNITY DETECTION ALGORITHMS ON DIFFERENT REAL NETWORKS

PARITA OZA, SMITA AGRAWAL, DHRUV RAVALIYA, AND RIYA KAKKAR*

Abstract. Complex networks are an essential tool in machine learning and data mining. The underlying information can help understand the system and reveal new information. Community is sub-groups in networks that are densely connected. This community can help us reveal a lot of information. The community detection problem is a method to find communities in the network. The igraph library is used by many researchers due to the utilization of various community detection algorithms implemented in both Python and R language. The algorithms are implemented using various methods showing various performance results. We have evaluated the community detection algorithm and ranked it based on its performance in different scenarios and various performance metrics. The results show that the Multi-level, Leiden community detection algorithm, and Walk trap got the highest performance compared to spin glass and leading eigenvector algorithms. The findings based on these algorithms help researchers to choose algorithms from the igraph library according to their requirements.

Key words: Community detection, igraph, Multi-level, Walk trap, Leiden

1. Introduction.

In computer science, network theory is a remarkable field that uses various techniques to utilize the vast scale of graphs consisting of numerous nodes and patterns connected in a network. Network theory helps in modeling computer networks, biological neural networks, traffic networks, protein networks, etc. One significant feature of these networks is community structure, i.e., densely connected components networks. The most crucial task in network analysis involves detecting communities in a network. We may have millions of nodes and edges in a colossal network connected in a quiet complex manner [17]. For example, nowadays, people use various online social networks such as Twitter, Facebook, and Instagram, which connect massive numbers of people forming a complex network [11]. So, it becomes a tremendous effort to detect communities in such networks. For that, many research work worldwide have been discussed and compared considering the community detection algorithms in which each of them associates various techniques for community detection [3] [18] [4]. Many research works have utilized machine learning [15] and deep learning networks [16] in community detection. For example, artificial networks [13] [14] or generating benchmark networks [10] can be considered with their implementation of various algorithms. However, most of the research works did not consider the performance, modularity, and coverage of the network [24] [5]. Therefore, we have evaluated the different algorithms implemented in the igraph library on different sizes of real networks. igraph is a set of tools for generating, altering, and analyzing graphs and networks. The techniques are commonly used in academic network science research. Further, we have applied the igraph library algorithms on real networks and ranked them based on their performance measures. We present the comparison of various igraph community detection algorithms considering the factors such as modularity, coverage, and performance to evaluate the impact of the algorithms on the network. We have computed the final rank considering the mean rank of all the possible community detection algorithms to evaluate the performance of the network. In nutshell, Our main research objective is to present a comparative analysis of the various algorithm on different datasets of sparse to dense networks and their applicability hence we utilized algorithms pertaining to the same library.

1.1. Research Contributions.

- We present the analysis of different igraph community detection algorithms such as Multi-level, Leiden, and Walk trap on different real networks.
- We discuss the various community detection algorithms implemented in the igraph library.

Department of Computer Science and Engineering, Nirma University, Ahmedabad, India (18bce195@nirmauni.ac.in, 21ftphde56@nirmauni.ac.in, parita.prajapati@nirmauni.ac.in,smita.agrawal@nirmauni.ac.in)

Table 1.1: Abbreviations for the community detection algorithms

EB	Edge Betweenness Centrality Measure
LE	Leading Eigenvector
LP	Label Propagation
IM	Map of information
SG	Spin-Glass
LC	Leiden algorithm based on optimization of modularity
FG	Fast-Greedy
ML	Multi-level Modularity Optimization
WT	Walk-Trap

- We present a comparative study of the various algorithm on different dataset of sparse to dense network and it's applicability .
- Finally, the performance results have been analyzed considering the modularity, performance measure, and coverage of community detection algorithms.

1.2. Paper Organization. The rest of the paper is structured as follows: Section 2 presents summary of the evaluated algorithms. In section 3 we discuss various datasets utilized in this work. Section 4 presents the methodology we adopted for evaluation. Section 5 presents the result and discussion. We finally end with conclusion in section 6.

2. The evaluated algorithms. The igraph library implements various community detection algorithms, which are discussed in the next section. In the algorithms, V is considered the set of vertices/nodes in the network, and E is considered the set of edges.

Edge betweenness centrality measure. This algorithm is called edge_betweenness in the igraph library. It can work on directed and weighted edges and can also handle multiple components in the network. It was initially proposed by Girvan and Newman in 2002 [8]. Later, in 2004, they presented a new version using the modularity measure, which is a method that estimates how modular network partitions in a given network. The main idea is to calculate the betweenness centrality for all edges and gradually remove the ones with the highest value to form a dendrogram [8]. The time complexity of the algorithm is $O(VE^2)$.

Leading Eigenvector. This algorithm is called the leading eigenvector and it can handle multiple components. This algorithm was proposed by Newman *et al.* [12] in 2006 to find optimal modularity using eigenvectors. The graph is separated into two segments in each step so that the separation generates a considerable rise in modularity. The split is decided by assessing the modularity matrix leading to the eigenvector and a halting condition that prohibits tightly related groups from being divided further. The complexity of algorithm can be computed as $O(V^2 + E)$.

Label Propagation. In the igraph library, this method is known as label-propagation and works with weighted edges. Raghavan *et al.* [20] first proposed it in 2007, in which every node in this algorithm is given a unique label at the start. Further, the algorithm involves iteration, re-assigning labels to nodes so that each node receives the label that its neighbor most frequently uses. The procedure ends when each node's label is one of the most common labels in its neighborhood. Communities are connected subgroups of nodes with the same labeled neighbors. The complexity of the algorithm can be calculated as $O(E + V)$.

Map of information. This algorithm is known as InfoMAP in the igraph library and works on weighted and directed edges. In 2008, Rosvall *et al.* [22] proposed the algorithm based on information-theoretic principles. With the help of random walks, it builds a grouping that provides the shortest description length for a random walk on the graph. The description length is measured by the expected number of bits per vertex required to encode the path of a random walk[8]. The complexity of the algorithm is $O(V * (E + V))$.

Spin Glass. In the igraph library, this algorithm is known as spin glass, and it can function with both directed and weighted edges and numerous components. Reichardt *et al.* [21] proposed this method in 2006 by introducing the technique based on the Potts model. In this model, each particle/vertex can be in one of the multiple spin states, and the interactions between the particles/edges decide which vertices prefer to be in

Table 3.1: Summary of datasets

Database –>	**Zachary's Karate Club**	**Football Network**	**HEP-TH Network**
Nodes	34	35	8400
Edges	78	118	15800
Density	0.139037	0.198319	0.000450686
Maximum degree	17	19	50
Minimum degree	1	1	0
Average degree	4	6	3
Average clustering coefficient	0.570638	0.338986	0.441964
Number of triangles	135	351	39900

the same spin state and which prefer to be in other spin states. The spin configuration reduces the spin glass's energy, and further spin states of the particles define the communities at the end.

Leiden algorithm based on optimization of modularity. The igraph library name for this algorithm is Leiden, and it works on both weighted and unweighted edges. In 2019, Traag *et al.* [23] proposed the Leiden optimization algorithm, which resembles the Multi-level algorithm in terms of functioning, except it is faster and yields better results. It can help with modularity and the Constant Potts Model, unaffected by the resolution constraint. If we consider the sparse graph for computing the complexity of the algorithm, it is found to be linear.

Fast-Greedy. This algorithm is known as fast greedy in the igraph library and it can work on weighted edges and handle multiple components. In 2004, Newman proposed this algorithm for detecting communities through a hierarchical approach rather than betweenness centrality; it is completely based on the modularity measure [7]. Later, Clauset *et al.* [7] suggested an optimized greedy version of Newman's proposed algorithm, which uses a more efficient data structure and an improved version of the Modularity measure. The complexity of algorithm can be determined as $O(VElogV)$.

Multi-level Modularity Optimization. This algorithm is termed multi-level modularity optimization in the igraph library and works on the weighted edges. This algorithm was proposed by Blondel *et al.* [6] in 2008, which resembles the fast greedy algorithm, but at each level of the dendogram, it considers local modularity, where a node is merged with the local neighbor to achieve the highest contribution to modularity. The complexity of algorithm is linear and can be calculated as $O(E)$.

Walk-Trap. This type of algorithm can work with weighted edges. Pons *et al.* [19] reportedly introduced it in 2006 to utilize the features of random walks. Moreover, there is a high probability of random walks in the same community. The complexity of the algorithm can be determined as $O(EV^2)$.

3. Datasets. We utilized three datasets in our work; the Karate network, the football network, and the Hep-Th network. This dataset represents a real network scenario. Zachary's Karate Club is being used for more minor test cases. It is based on Zachary's 1977 model of a friendship network between 34 karate club members at a US university with 78 interrelationships [25], which can be referred to as the Karate Network. The football network [8] is the American football college team dataset and has medium larger test cases. The dataset, HEP-TH (high energy physics-Theory) [9] is a large network consisting of details of authors and related papers submitted to the high energy physics-Theory category. We present a summary of this dataset in table 3.1.

4. Methodology for Evaluation. The metrics considered for evaluation are modularity, time, and coverage. The modularity measurement Q and the running time t returned by the algorithms are used for the evaluation. Alternatively, they can be arranged based on the Q/t ratio. According to the different running times of algorithms, a sigmoid function is being used to rescale the t values to a smaller interval and enhance the weight of the modularity metric in the ranking process. The above associations can be represented as follows:

$$S(x) = \frac{1}{(1 + e^{-x})} \tag{4.1}$$

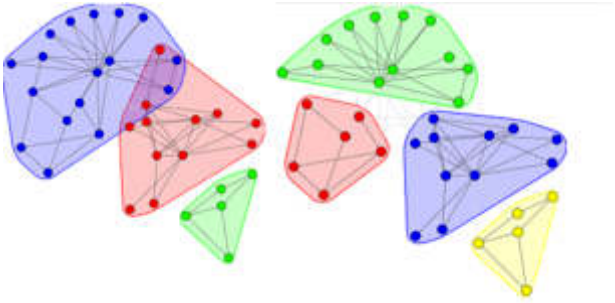


Fig. 4.1: Communities detected by the Leiden and Multi-level detection algorithm for Karate Network.

Table 4.1: Different igraph's community detection algorithm performance and results for the Karate Network.

Algorithm	**C**	**Q**	**T**	**P**	**Rank**
EB	5	0.40	1.99 ms	0.80	4
LE	4	0.39	6.98ms	0.78	5
LP	2	0.35	991µs	0.71	8
IM	3	0.40	6.98 ms	0.80	3
SG	4	0.41	190ms	0.76	6
LC	3	0.40	996µs	0.80	2
FG	3	0.38	998µs	0.76	7
ML	4	0.41	999 µs	0.83	1
WT	5	0.35	995µs	0.70	9

The modularity Q can be calculated with the help of the igraph method, i.e., self graph modularity. It can be evaluated based on the performance measure P, which can be mentioned as follows:

$$P_k = \frac{Q_k}{Sig(t_k)} \tag{4.2}$$

where Q_k is the modularity measure and t_k is the running time of algorithm k.

Coverage can be considered to detect the communities close to the desirable communities, which can be calculated with the help of partitions. Each partition's coverage is defined by the ratio of inter-communities edges to the total number of edges in the graph. The intra-community edges are those formed by joining a pair of nodes in the same partition block. In an ideal community, each node is connected to every other node of all the edges of the graph within clusters that lead to the coverage of one [2, 1].

The final rank can be calculated considering the average of the partial ranks algorithms obtained across all the test cases. This helps us to compare and rank the overall results of the algorithms.

5. Result and Discussion. The simulation has been performed considering the various community detection algorithms along with their parameters. The results of the Karate Network have been analyzed as shown in Table 4.1 in which C denotes the amount of detected community. It can be observed that multi-level and infomap achieve the same modularity; however, multi-level community detection yields better performance. While the Label propagation algorithm achieves the lowest modularity and the spin glass algorithm achieves the highest modularity. Further Fig. 4.1 shows a visual representation of the community detected by the Multi-level and Leiden community detection algorithm for Karate Network. Then, Fig. 4.2 and Fig. 4.3 visualize the community detection by Multi-level and Leiden community detection algorithm for Football and Hep-Th Network, respectively.

In the next scenario, as shown in Table 4.2, different community detection algorithms have been analyzed

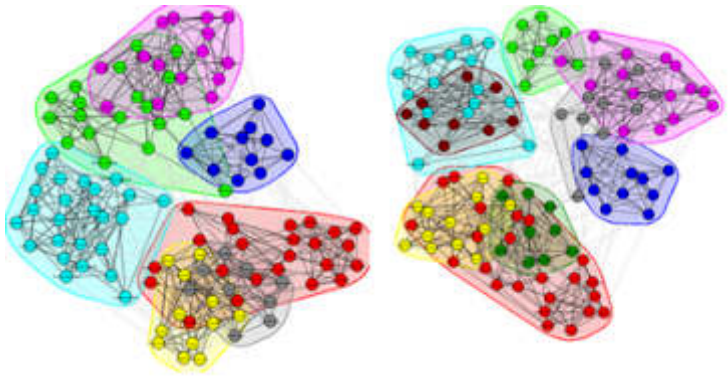


Fig. 4.2: Communities detected by the Leiden and Multi-level detection algorithm for Football Network

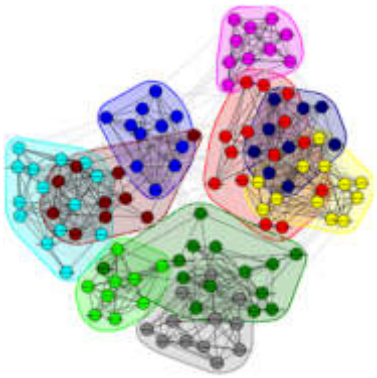


Fig. 4.3: Communities detected by the Leiden and Multi-level detection algorithm for Hep-Th Network

Table 4.2: Different igraph's community detection algorithm performance and results for Football Network.

Algorithm	**C**	**Q**	**t**	**P**	**Rank**
EB	10	0.60	234 ms	1.18	6
LE	8	0.48	23.9 ms	0.97	8
LP	10	0.60	987 µs	1.20	3
IM	13	0.55	18 ms	1.11	7
SG	11	0.60	736 ms	1.20	4
LC	11	0.60	998 µs	1.20	5
FG	-	-	-	-	9
ML	10	0.60	998 µs	1.21	1
WT	10	0.60	2 ms	1.20	2

for Football Network in which the multi-level detection algorithm yields the highest modularity leading to the highest rank. Compared to the previous scenario, Label propagation gives good results in scaling leading into a higher rank. In the final scenario, community detection algorithms have been analyzed for Hep-Th Network which is a large network comprising of various nodes and edges as shown in Table 4.3 in which multi-level detection algorithm results into improved modularity which results into the highest rank. However, it can effect the peformance of the network due to the high complexity of computation time as it contains large dataset of

Table 4.3: Different igraph's community detection algorithm performance and results for Hep-Th Network.

Algorithm	**C**	**Q**	**t**	**P**	**Rank**
EB	-	-	-	-	8
LE	1366	0.75	3.17 s	1.30	6
LP	1859	0.76	66.8 ms	1.52	4
IM	1847	0.76	6.46 s	1.16	7
SG	-	-	-	-	9
LC	2187	0.71	63.8 ms	1.42	5
FG	1411	0.81	151 ms	1.61	2
ML	1379	0.84	57.8 ms	1.69	1
WT	1411	0.81	703 ms	1.56	3

Table 4.4: Different igraph's community detection algorithm final rank on various network.

Algorithm	**Karate**	**Football**	**Hep-Th**	**Mean**	**Rank**
EB	4	6	8	6	6
LE	5	8	6	6.33	8
LP	8	3	4	5	4
IM	3	7	7	5.67	5
SG	6	4	9	6.33	9
LC	2	5	5	4	2
FG	7	9	2	6	7
ML	1	1	1	1	1
WT	9	2	3	4.677	3

Table 4.5: Different igraph's community detection algorithm coverage on various datasets

Algorithm	**Karate Network**	**Football Network**	**Hep-Th Network**
EB	0.69	0.71	0
LE	0.66	0.63	0.82
LP	0.85	0.71	0.78
IM	0.82	0.63	0.77
SG	0.73	0.69	0
LC	0.82	0.60	0.71
FG	0.75	0	0.90
ML	0.73	0.71	0.88
WT	0.58	0.70	0.85

networks which increases the overall complexity of the network. Label propagation detection continues to be good at scaling on the larger dataset. Similarly, Multi-level and Fast greedy seems to provide more modularity to the network improving the performance of the network. Nevertheless, we have considered small and large dataset to evaluate the performance of the various community detection algorithms considering Karate, Football, and Hep-Th network.

The results can be estimated by evaluating the final rank for all the community detection algorithms. Table 4.4 shows the comparison of final rank evaluated by considering the mean of rank in all the possible scenarios discussed. Further, Table 4.5 presents the coverage measure of various community detection algorithms. For the smaller dataset, Label propagation, Infomap and Ledian algorithm has one of highest results for denser

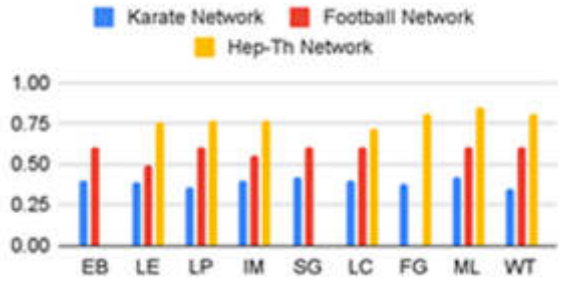


Fig. 4.4: Modularity of different algorithms

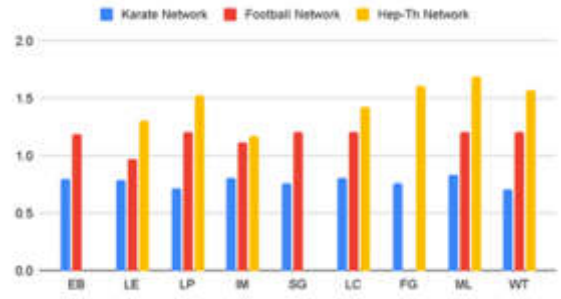


Fig. 4.5: Performance measure of the different algorithms

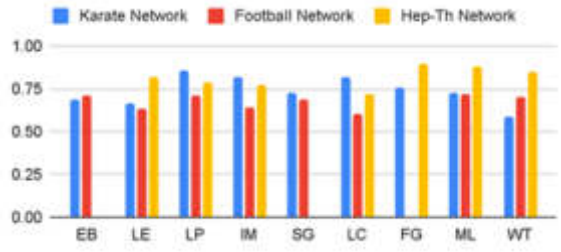


Fig. 4.6: Coverage of the different algorithms

partitions. While Fast greedy, Multi-level and Walktrap performs better for larger dataset. On the other hand,figures 4.4, 4.5 and 4.6 depicts the modularity, performance measure, and coverage of different algorithms for community detection.

6. Conclusion. In this paper, We have considered the community detection algorithms and estimated performance-based evaluation for their implementation in the igraph library. The evaluation was done based on their strength to detect communities in different types and sizes of datasets. The performance metrics to evaluate these algorithms used are modularity, performance measure, and coverage. Further, the algorithms are ranked based on performance measures. Results indicate that Multi-level performed best, followed by Leiden and Walk trap, which achieved lower performance in the smaller test case. Therefore, the evaluation of various igraph community detection algorithms has been estimated on the real network to show the impact of modularity, coverage, and performance measure.

REFERENCES

[1] S. Agrawal and A. Patel, *Csg cluster: A collaborative similarity based graph clustering for community detection in complex networks*, International Journal of Engineering and Advanced Technology, 8 (2019), pp. 1682–1687.

[2] S. Agrawal and A. Patel, *Clustering algorithm for community detection in complex network: a comprehensive review*, Recent Advances in Computer Science and Communications (Formerly: Recent Patents on Computer Science), 13 (2020), pp. 542–549.

[3] S. Agrawal and A. Patel, *Sag cluster: An unsupervised graph clustering based on collaborative similarity for community detection in complex networks*, Physica A: Statistical Mechanics and its Applications, 563 (2021), p. 125459.

[4] S. E. Amrahov and B. Tugrul, *A community detection algorithm on graph data*, in 2018 International Conference on Artificial Intelligence and Data Processing (IDAP), 2018, pp. 1–4.

[5] F. Berardo de Sousa and L. Zhao, *Evaluating and comparing the igraph community detection algorithms*, in 2014 Brazilian Conference on Intelligent Systems, 2014, pp. 408–413.

[6] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, *Fast unfolding of communities in large networks*, Journal of statistical mechanics: theory and experiment, 2008 (2008), p. P10008.

[7] A. Clauset, M. E. J. Newman, and C. Moore, *Finding community structure in very large networks*, Phys. Rev. E, 70 (2004), p. 066111.

[8] M. Girvan and M. E. Newman, *Community structure in social and biological networks*, Proceedings of the national academy of sciences, 99 (2002), pp. 7821–7826.

[9] J. Leskovec, J. Kleinberg, and C. Faloutsos, *Graph evolution: Densification and shrinking diameters*, ACM transactions on Knowledge Discovery from Data (TKDD), 1 (2007), pp. 2–es.

[10] J. Leskovec, K. J. Lang, and M. Mahoney, *Empirical comparison of algorithms for network community detection*, in Proceedings of the 19th international conference on World wide web, 2010, pp. 631–640.

[11] H. S. Nalawala, J. Shah, S. Agrawal, and P. Oza, *A comprehensive study of "etcd"—an open-source distributed key-value store with relevant distributed databases*, in Emerging Technologies for Computing, Communication and Smart Cities: Proceedings of ETCCS 2021, Springer, 2022, pp. 481–489.

[12] M. E. J. Newman, *Finding community structure in networks using the eigenvectors of matrices*, Phys. Rev. E, 74 (2006), p. 036104.

[13] G. K. Orman and V. Labatut, *A comparison of community detection algorithms on artificial networks*, in International conference on discovery science, Springer, 2009, pp. 242–256.

[14] G. K. Orman, V. Labatut, and H. Cherifi, *Qualitative comparison of community detection algorithms*, in International conference on digital information and communication technology and its applications, Springer, 2011, pp. 265–279.

[15] P. Oza, P. Sharma, and S. Patel, *Machine learning applications for computer-aided medical diagnostics*, in Proceedings of Second International Conference on Computing, Communications, and Cyber-Security: IC4S 2020, Springer, 2021, pp. 377–392.

[16] ———, *A drive through computer-aided diagnosis of breast cancer: a comprehensive study of clinical and technical aspects*, Recent Innovations in Computing: Proceedings of ICRIC 2021, Volume 1, (2022), pp. 233–249.

[17] A. Patel, P. Oza, and S. Agrawal, *Sentiment analysis of customer feedback and reviews for airline services using language representation model*, Procedia Computer Science, 218 (2023), pp. 2459–2467.

[18] N. Patel, P. Oza, and S. Agrawal, *Homomorphic cryptography and its applications in various domains*, in International Conference on Innovative Computing and Communications, S. Bhattacharyya, A. E. Hassanien, D. Gupta, A. Khanna, and I. Pan, eds., Singapore, 2019, Springer Singapore, pp. 269–278.

[19] P. Pons and M. Latapy, *Computing communities in large networks using random walks*, in International symposium on computer and information sciences, Springer, 2005, pp. 284–293.

[20] U. N. Raghavan, R. Albert, and S. Kumara, *Near linear time algorithm to detect community structures in large-scale networks*, Physical review E, 76 (2007), p. 036106.

[21] J. Reichardt and S. Bornholdt, *Statistical mechanics of community detection*, Phys. Rev. E, 74 (2006), p. 016110.

[22] M. Rosvall and C. T. Bergstrom, *Maps of information flow reveal community structure in complex networks*, arXiv preprint physics.soc-ph/0707.0609, (2007).

[23] V. A. Traag, L. Waltman, and N. J. Van Eck, *From louvain to leiden: guaranteeing well-connected communities*, Scientific reports, 9 (2019), pp. 1–12.

[24] N. Xiao, *Community detection and key nodes of complex technology exchange network*, in 2016 2nd International Conference on Cloud Computing and Internet of Things (CCIOT), 2016, pp. 101–104.

[25] W. W. Zachary, *An information flow model for conflict and fission in small groups*, Journal of anthropological research, 33 (1977), pp. 452–473.

Edited by: Katarzyna Wasielewska
Received: Feb 16, 2023
Accepted: Jul 9, 2023


Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2023 SCPE. Volume 24, Issues 2, pp. 181–190, DOI 10.12694/scpe.v24i2.2238

RESEARCH HIGHLIGHT GENERATION WITH ELMO CONTEXTUAL EMBEDDINGS

TOHIDA REHMAN*, DEBARSHI KUMAR SANYAL†, AND SAMIRAN CHATTOPADHYAY‡

Abstract. With the advent of digital publishing and online databases, the volume of textual data generated by scientific research has increased exponentially. This makes it increasingly difficult for academics to keep up with new breakthroughs and synthesise important information for their own work. Abstracts have long been a standard feature of scientific papers, providing a concise summary of the paper's content and main findings. In recent years, some journals have begun to provide research highlights as an additional summary of the paper. The aim of this article is to create research highlights automatically by using various sections of a research paper as input. We employ a pointer-generator network with a coverage mechanism and pretrained ELMo contextual embeddings to generate the highlights. Our experiments shows that the proposed model outperforms several competitive models in the literature in terms of ROUGE, METEOR, BERTScore, and MoverScore metrics.

Key words: deep learning, pointer-generator network, ELMo, natural language generation.

AMS subject classifications. 68T07, 68T50

1. Introduction.

Van Noorden et al.[1] estimated that every nine years the quantity of scientific articles roughly doubles. While the growth in the number of scientific papers is generally viewed as a positive development, it also presents some challenges. With the exponential growth in the number of scientific papers being published [2], it can be challenging for researchers to stay up-to-date with the latest findings in their field and identify which papers are the most important. Although both research highlights and abstracts are summaries of the research article, they serve different objectives and have distinct qualities. The research highlights typically consist of a brief, bulleted list of the main findings of the paper. By providing a concise summary of the most important aspects of the research, the highlights can help readers to quickly assess the relevance of the paper and determine whether it is worth reading in full. The main findings and contributions of the paper can be emphasized in promotional materials such as social media posts by utilizing the highlights.

Text summarization is a technique to prepare a compact text document that has been condensed while retaining its most important and salient information. Text summarization is classified into two types [3]: extractive summarization, which involves picking and combining relevant sentences or phrases directly from the original text [4], but abstractive summarization, which involves generation and formation of new sentences which capture the essence of the original text [5]. In this paper, we propose a method to generate research highlights (a form of summary) of a research paper using deep neural network-based model. Unlike large pretrained language models (often called foundation models [6]) that require access to a huge document corpus, large training time, a huge energy expenditure, our proposed method is task-specific, utilizes pretrained embeddings, and is trained with a much smaller domain-specific corpus. Thus, it is a scalable model suitable for generation of research highlights from scientific papers. The primary contributions of this article are:

1. Our proposition involves integrating ELMo embeddings with pointer-generator networks that utilize coverage mechanisms.
2. We examine how well the proposed model can generate research highlights using two different types of inputs: (a) only the abstract, and (b) a combination of the abstract, introduction, and conclusion of a

*Department of Information Technology, Jadavpur University Salt Lake Campus, Kolkata-700106, West Bengal, India. (tohidarehman.it@jadavpuruniversity.in)

†School of Mathematical & Computational Sciences, Indian Association for the Cultivation of Science, Jadavpur, Kolkata-700032, West Bengal, India. (debarshi.sanyal@iacs.res.in)

‡Department of Information Technology, Jadavpur University, Salt Lake Campus, Kolkata-700106, West Bengal, India. (samiran.chattopadhyay@jadavpuruniversity.in)

research paper.

3. We evaluate our models extensively through multiple metrics, including ROUGE [7], METEOR [8], BERTScore [9], and MoverScore [10] metrics. We show that the proposed model outperforms other existing techniques available in the literature. We also identify the role of each component of our model using an ablation study.

2. Literature survey. Extractive approaches are a text summarization technique that focuses on identifying the most important phrases or sentences from the source text and present them as a summary. Luhn et al.[3] proposed a method of text summarization to select high-scored sentences based on the high frequency words while ignoring the common words. Baxendale et al. [11] proposed using the position of a sentence to select the important sentences of a document. In that research found that 85% of the theme sentences selected from the first sentences of the paragraph and 7% as the last sentence of a paragraphs. Edmundson et al.[12] proposed a method for automatically summarising texts that assigns a score to each sentence based on four features including sentence position, word frequency, document skeleton, and cue words. Kavita et al. [13] proposed a graph based abstractive model called "Opinosis" useful for highly repetitive opinions. The progress of sequence-to-sequence (Seq2Seq) models has significantly improved the state-of-the-art in abstractive summarization [14]. A neural network-based Seq2Seq models used to learn a map of a sequence of input tokens to a sequence of output tokens. Bahdanau et al. [15] improved the fundamental encoder and decoder models' performance. Chopra et al. [16] proposed a unique "Convolutional Attention-based Conditional Recurrent Neural Network (CARCNN)" architecture for abstractive text summarization. On the Gigaword Corpus and the DUC 2004 datasets, the proposed model was evaluated. Nallapati et al. [5] proposed an abstractive text summarization technique that uses "Attentional Encoder-Decoder Recurrent Neural Networks". The authors proposed a model that leverages this architecture to generate a summary of a given input document. Using bidirectional recurrent neural network the model first encodes the input document, which captures the input's contextual information. At the docedoer end, the summary is then generated one word at a time by considering encoded inputted document. The attention mechanism helps the model to concentrate on the key passages of the text when generating the summary. See et al.[17] proposed a model to overcome the problem of out-of-vocabulary words (OOV) and repetition words generation named as pointer-generator network with coverage mechanism. Coverage mechanism helps to avoid repetition by keeping track of what has been summarized by pointing and copy words from the inputted text. Gehrmann et al.[18] proposed a model "Bottom-Up Abstractive Summarization" which enhances the capacity to condense content, while still creating fluent summaries. Liu et al. [19] proposed a transformer-based model for abstractive summary generation of Wikipedia articles. For better word semantic representation, a model combining the pointer-generator model with two pre-trained word embeddings—word2vec and FastText [20].

Scientific paper summarization can be broadly categorized into two types: abstract generation from the paper and summary generation based on the citation [21]. In the past, extractive summarization methods have been widely used for summarizing scientific articles. Kupiec et al.[4] used a limited dataset of 188 scientific document and summary pairs. This model used a set of features to rank sentences for scientific paper summarization. Contractor et al. [22] proposed a model for extractive summarization to utilize the concept of argumentative zones (AZs) framework for academic papers. Kinugawa and Tsuruoka [23] proposed a two-level hierarchical structure based on encoder-decoder for extractive summarization of research papers.

But a common trend nowadays to supplement research highlights with full research paper and abstract. Hence, Highlights generation is another categorization task of text summarization. Collins et al. [24] proposed a supervised extarctive model for identifying a sentence is highlights or not also published urls for computer science publications as a benchmark dataset named CSPubSum. Alambo et al. [25] proposed a techniques for selecting salient language units and producing text in order to produce an abstractive summary of a scientific paper. L. Cagliero et al. [26] proposed an extractive approach based on gradient boosting method to select some sentences as a research highligths. Rehman at el. [27] proposed an abstractive method to generate research highlights from a research paper's abstract, by combining a pointer-generator model with Glove embeddings.

Our work is significantly different from the above works. Here we use pretrained ELMo embeddings with a pointer-generator model with coverage mechanism and generate research highlights using various sections of the paper. Generic texts are used to train the pretrained models like PEGASUS [28], T5 [29], (GPT(Generative

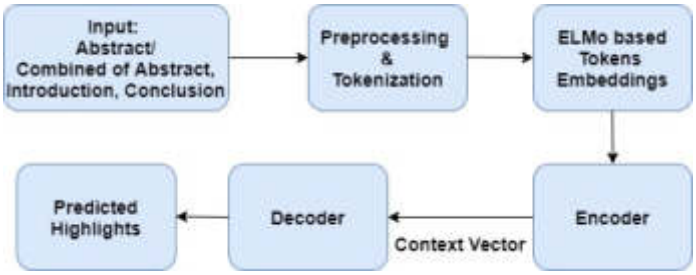


Fig. 3.1: Workflow used in our encoder-decoder model.

Pre-trained Transformer)-like) decoder [30], and BART [31] which appear to be the most effective summarizer. In order to have even better performance, they need to be fine-tuned using domain-specific inputs. However, this process needs significant resources. Our proposal describes a cost effective but useful architecture to meet the same objective.

3. Methodology. In this section, we describe the model we use to generate research highlights from the research articles. The workflow of our system is shown in Figure 3.1.

We conduct experiments with four different variations: (1) Pointer-generator model proposed by [17], (2) Incorporating coverage mechanism (proposed in [32]) into the pointer-generator model (the combined model is also referred to in the same work [17]), (3) Pretrained ELMo embeddings [33] with pointer-generator model, and (4) Pretrained ELMo embeddings with pointer-generator model and coverage mechanism.

3.1. ELMo Pre-trained word representations. Pre-trained word representations [34] have a significant role in many neural language understanding models. Pre-trained word representations are a key component in many neural language understanding models. On the other hand representation of learning with high quality and accurate generation are very challenging. A deep contextualized word representation, ELMo [33] can be used to capture the complex behaviour of word syntax, semantics and linguistic context. ELMo employs vectors that was trained with a stacked bidirectional LSTM.

3.2. Pointer-generator model with ELMo embeddings and coverage mechanism. We use a pointer-generator model, proposed by See, Liu and Manning [17]. It consists of a seq2seq model with a BiLSTM encoder and an LSTM decoder with attention [5]. However, instead of using word embeddings trained from scratch or non-contextual embeddings like word2vec [35] or GloVe [36], we use context-sensitive embeddings that represent homonyms with different vectors, despite the fact that these words have the same spelling. In other words, word representations capture the fine differences in meaning that arise from the context in which the words are used. In our present work, we add a pretrained ELMo contextual embeddings [33] layer that generates embedding for each word of the input text. Instead of directly passing the individual token id to the encoder recurrent neural network, we feed the token embeddings prepared by pretrained ELMo [33] embeddings layer. This can improve the model's ability to generate hidden states because the input words embedding matrix is initialized with the pretrained word embeddings ELMo. The embeddings are fine-tuned during model training. The dimension of ELMo word embeddings used in our experiment is 1024. The decoder has a unique copying technique, which decides between *copying* a word from the source text by utilizing copying mechanism or *generating* new words from the vocabulary (built from the vocabulary of the whole training corpus and the current input document). The copying mechanism helps to deal with out-of-vocabulary (OOV) words. The generating mechanism, on the other hand, induces new words which indicate novel paraphrasing. The decoder strikes a balance between copying words and generating words using a hyperparameter, which probabilistically chooses between the two alternatives. However, the pointer-generator model sometimes generates the same words repetitively. To overcome this problem, we used the coverage mechanism of Tu et al.[32]. In essence, this model focuses on the preceding time steps of the decoder through attention so that attending to the same word in the input document again and again is penalized.

4. Experimental setup.

4.1. Datasets. We make use of the computer science publication benchmark dataset published by Collins et al.[24] named CSPubSum, which contains the URLs of $\sim$ 10K papers from ScienceDirect [1]. The following fields are typically present in the documents: title, abstract, author-written research highlights, authors-written keywords, introduction, related work, experiment, conclusion, and other major subsections that are part of the discourse structure of a research paper. For our experiments, we divided the dataset into training, validation, and test subsets (train, val, test) in proportion of $80:10:10$. In 98% of the papers, the highlights are at least 1.5 times shorter than the abstract. Therefore, research highlights can be viewed as a summary of both the abstract and the paper.

4.2. Data pre-processing. Before inputting the dataset to the model, we did some basic pre-processing steps. We removed unintended symbols, letters, urls, HTML tags and special characters. Then we changed the dataset to lowercase. To conduct experiments, we arranged the dataset in several ways. In particular, we organized it as *(abstract, research highlights written by author)*, *(abstract* $\bigoplus$ *introduction* $\bigoplus$ *conclusion, research highlights written by author)*, where text concatenation is represented as'$\bigoplus$'. When considering only abstract as an input, we allowed a maximum of 400 tokens. In the case of combined inputs from abstract, introduction and conclusion sections, we allowed a maximum of 1500 tokens. From each section, we allowed up to 500 tokens. In all cases, the token count of model-generated research highlights was limited to 100 only.

4.3. Implementation details. All the models were trained on the GPU-supported Colab Pro+ environment. The pointer-generator network with ELMo embeddings used 1024 as the word embedding dimension and that without ELMo embeddings used 128 as the word embedding dimension. For all models, the maximum vocabulary size was restricted to 50K tokens. For all models, the dimension of RNN hidden states is 256. We chose maximum gradient norm of 1.2 for gradient clipping.

4.4. Evaluation metrics. To compare the performance of the various models, we used the following metrics: ROUGE [7], METEOR [8], BERTscore [9], and MoverScore [10]. We have used ROUGE [7] metric to measure the word overlap between the research highlights written by the authors (ARRHS) and those generated by model (MGRHS). The recall (R), precision (P) and F1-measure ($F1$) for ROUGE-N are calculated as follows:

$$R = \frac{Matched\ number\ of\ n/grams\ in\ (\texttt{MGRHS}, \texttt{ARRHS})}{Number\ of\ n/grams\ in\ \texttt{ARRHS}} \tag{4.1}$$

$$P = \frac{Matched\ number\ of\ n/grams\ in\ (\texttt{MGRHS}, \texttt{ARRHS})}{Number\ of\ n/grams\ in\ \texttt{MGRHS}} \tag{4.2}$$

$$F1 = \frac{2 * (R * P)}{R + P} \tag{4.3}$$

A sequence of n words makes up an n-gram. ROUGE-L measures the longest common subsequence (LCS) between MGRHS and ARRHS. ROUGE-S measures the skip-bigram matched between MGRHS and ARRHS where a skip-bigram is a bigram that allows random word-gaps or skips between words. The recall based on skip-bigram is calculated as follows:

$$R_{skip} = \frac{Matched\ number\ of\ Skip/bigrams\ in\ (\texttt{MGRHS}, \texttt{ARRHS})}{Number\ of\ Skip/bigrams\ in\ \texttt{ARRHS}} \tag{4.4}$$

The precision based on skip-bigram is calculated as follows:

$$P_{skip} = \frac{Matched\ number\ of\ Skip/bigrams\ in\ (\texttt{MGRHS}, \texttt{ARRHS})}{Number\ of\ Skip/bigrams\ in\ \texttt{MGRHS}} \tag{4.5}$$

[1] (https://www.sciencedirect.com)

The F1-measure ($F1_{skip}$) based on skip-bigram is calculated as follows:

$$F1_{skip} = \frac{2 * (R_{skip} * P_{skip})}{R_{skip} + P_{skip}} \tag{4.6}$$

ROUGE-SU is an extension of ROUGE-S which counts both skip-bigram and unigram between MGRHS and ARRHS. According to the official ROUGE script, all of our ROUGE scores have a 95% confidence interval of at most ±0.25.

METEOR-score is calculated using an explicit word-to-word correspondence of research highlights generated by the model (MGRHS) and research highlights written by the authors (ARRHS).

BERTScore is calculated based on pairwise cosine similarity of each token in the research highlights generated by model (MGRHS) with that in the highlights written by authors (ARRHS). Here, instead of using the tokens directly, the similarity is computed based on contextual embeddings. When we tokenize the research highlights written by authors (ARRHS) and pass the tokens through the embedding model (in our case, ELMo), we get a sequence of contextual embeddings denoted as $\vec{x} = \langle \vec{x}_1, \ldots, \vec{x}_m \rangle$. Similarly, when we tokenize the research highlights generated by the model (MGRHS) and embed the tokens, we get a sequence of contextual embeddings denoted as $\hat{\vec{x}} = \langle \hat{\vec{x}}_1, \ldots, \hat{\vec{x}}_n \rangle$. The values of recall ($R_{\text{BERT}}$), precision ($P_{\text{BERT}}$), and F1-scores ($F_{\text{BERT}}$) are computed as follows:

$$R_{\text{BERT}} = \frac{1}{m} \sum_{\vec{x}_i \in \vec{x}} \max_{\hat{\vec{x}}_j \in \hat{\vec{x}}} \vec{x}_i^{\top} \hat{\vec{x}}_j \qquad P_{\text{BERT}} = \frac{1}{n} \sum_{\hat{\vec{x}}_j \in \hat{\vec{x}}} \max_{\vec{x}_i \in \vec{x}} \vec{x}_i^{\top} \hat{\vec{x}}_j \tag{4.7}$$

$$F_{\text{BERT}} = \frac{2 * (R_{\text{BERT}} * P_{\text{BERT}})}{R_{\text{BERT}} + P_{\text{BERT}}} \tag{4.8}$$

MoverScore [10] is calculated based on the contextualized representations and Word Mover's Distance (WMD) [37] between the research highlights generated by model (MGRHS) and the research highlights written by authors (ARRHS). It can take into account the presence of new or unseen words in the generated text, and evaluate how well they fit into the overall structure and content of the original text. It allows many-to-one alignment to map the semantically similar words in MGRHS and ARRHS whereas BERTScore considers only one-to-one alignment. The sentences of the research highlights written by the authors (ARRHS) and the research highlights generated by the model (MGRHS) are represented as x and $\hat{x}$. Their sequence of n-grams are denoted as x^n and $\hat{x}^n$. The transportation cost matrix (C) is calculated based on a distance metric (d) between the n-grams as follows:

$$C_{i,j} = d(x_i^n, \hat{x}_j^n) \tag{4.9}$$

where $d(x_i^n, \hat{x}_j^n)$ is the Euclidean distance between the i-th n-gram of x and the j-th n-gram of $\hat{x}$ where both the n-grams are represented by their respective embeddings. The authors in [10] define a transportation flow matrix F where $F(i, j)$ captures the amount of flow from the i-th n-gram (x_i^n) in x^n to the j-th n-gram ($\hat{x}_j^n$) in $\hat{x}^n$. Let $\langle C, F \rangle$ denote the sum of all elements in the matrix obtained from element-wise multiplication of C and F. We associate weights f_{x^n} and $f_{\hat{x}^n}$ with the n-grams x_n and $\hat{x}_n$, such that each n-gram gets a single weight value in each case and assume that each of f_{x^n} and $f_{\hat{x}^n}$ defines a probability distribution (i.e., the entries of each vector sums to 1). Finally, the moverscore [10] is defined as

$$\text{WMD}(x^n, \hat{x}^n) = \min_{F \in \mathbf{R}^{|x^n| \times |\hat{x}^n|}} \langle C, F \rangle \qquad \text{such that } \; F\,\mathbf{1} = f_{x^n} \text{ and } \; F^{\top}\mathbf{1} = f_{\hat{x}^n} \tag{4.10}$$

5. Results.

5.1. Comparison of four pointer-generator model variants. In this section, we compare various scores predicted by four variants of model with different types of input cases. The four variants of the model are (1) Pointer-generation model (**PGM**), (2) Pointer-generation model with coverage (**PGM + Cov**), (3)

Table 5.1: Pointer-generator type model evaluation: ROUGE, METEOR, BERTScore and MoverScore scores on different inputs from the CSPubSum dataset

Input	Model Name	ROUGE-1	ROUGE-2	ROUGE-L	ROUGE-S	ROUGE-SU	METEOR	BERTScore	MoverScore
abstract only	PGM	35.44	11.57	29.88	11.45	12.35	25.4	83.80	56.69
	PGM + Cov	36.57	12.3	30.69	12.14	13.04	25.4	84.05	57
	PGM + ELMo	35.74	12.54	32.47	11.81	12.71	19.75	82.62	54.98
	PGM + ELMo + Cov	**38.4**	**13.32**	**35.45**	**13.41**	**14.35**	**30.61**	**86.65**	**57.94**
abstract +introduction +conclusion	PGM	33.49	10.83	30.87	10.67	11.59	25.51	86.01	56.75
	PGM + Cov	35.73	11.61	32.96	11.6	12.52	27.71	86.26	57.39
	PGM + ELMo	33.6	11.44	30.97	10.78	11.69	25.68	86.02	56.79
	PGM + ELMo + Cov	**36.34**	**12.11**	**33.77**	**11.98**	**12.97**	**27.78**	**86.68**	**57.63**

Table 5.2: Comparison of the performance of the proposed model with that of other approaches for CSPubSum data set.

Model Name	ROUGE-2 (F1)	ROUGE-L (F1)
LSTM Classification [24]	12.7	29.50
Gradient Boosting Regressor [26]	**13.9**	31.60
Pointer-generator+ Coverage + GloVe [27]	8.57	29.14
PGM + ELMo + Coverage	13.32	**35.45**

Pointer-generation model with ELMo embeddings (**PGM + ELMo**), (4) Pointer-generation model with ELMo embeddings and coverage mechanism (**PGM + ELMo + Cov**). In models (1) and (2), the input contains word embeddings that are randomly initialized and trained with the model. Since we have proposed model (4) in this paper, investigation of the other three variants may be seen as an ablation study. For each model, the input could be the abstract only or a combination of abstract, introduction and conclusion of the paper.

Input: Abstract only:

When the input is a research paper's abstract, the results for ROUGE-1, ROUGE-2, ROUGE-L, ROUGE-S, ROUGE-SU, METEOR, BERTScore and MoverScore are shown in Table 5.1. The pointer-generator model with ELMo embeddings and coverage mechanism achieves the best result in all cases.

Input: Abstract $\bigoplus$ Introduction $\bigoplus$ Conclusion:

When the input is a combination of abstract $\bigoplus$ introduction $\bigoplus$ conclusion, the results for ROUGE-1, ROUGE-2, ROUGE-L, ROUGE-S, ROUGE-SU, METEOR, BERTScore and MoverScore are shown in Table 5.1 The pointer-generator model with ELMo embeddings and coverage mechanism displays the highest performance in all cases.

5.2. Comparison with previous works. On the CSPubSum dataset, as shown in the Table 5.2, we compare the performance of our proposed work with other prior works. We notice that our model, pointer-generator with ELMo embeddings and coverage mechanism (**PGM + ELMo + Cov**) achieves higher ROUGE-L scores than other methods in the literature.

6. Case studies. In this section, we have shown a few research highlights generated by our models to enable a qualitative study of the performance of the models. In all case studies, yellow color represents factual errors, orange color represents repeating phrases and green color identifies some correctly added words or phrases.

Figure 6.1 shows the comparison of predicted research highlights generated by variants of pointer-generator models when the input (for training and test) is the abstract of a research paper. Observe that the first model **PGM** generates repeating phrase 'total route duration', which is solved when we add coverage mechanism. Similarly, a full sentence is repeated by the **PGM + ELMo** model which the fourth model **PGM + ELMo + Cov** corrects using the coverage mechanism. Note that the words 'subtour elimination constraints' that the last model (**PGM + ELMo + Cov**) generates is present in the abstract of the paper, and its insertion in the output is semantically correct, although it is absent in the golden set of research highlights submitted by the

Abstract: "This paper introduces a version of the classical traveling salesman problem with time-dependent service times. In our setting, the duration required to provide service to any customer is not fixed but defined as a function of the time at which service starts at that location. The objective is to minimize the total route duration, which consists of the total travel time plus the total service time. The proposed model can handle several types of service time functions, e.g., linear and quadratic functions. We describe basic properties for certain classes of service time functions, followed by the computation of valid lower and upper bounds. We apply several classes of subtour elimination constraints and measure their effect on the performance of our model. Numerical results obtained by implementing different linear and quadratic service time functions on several test instances are presented."
Author-written research highlights: "We introduce and model the traveling salesman problem with time dependent service times. The proposed model can handle several types of service time functions such as linear and quadratic. We describe properties of the service time function and compute valid lower and upper bounds. We propose and compare several mathematical programming formulations for the problem. Comparative computational results indicate that a flow based formulation is superior the other three."
PGM: "We study minimize the total route duration of the total route duration . Consider the total route duration of the total time [UNK] describe basic properties for certain classes of service time functions. Model can handle several types of service time functions e.g. linear and quadratic functions. Examples obtained by valid lower and upper bounds are presented. Results obtained by implementing different quadratic service time functions are discussed."
PGM + Cov: "We study the route duration of the total route duration of total service time. New model is based on various linear and valid lower heuristics. Results are conducted to illustrate different linear and quadratic service time functions. Results obtained by implementing linear and quadratic functions are presented."
PGM + ELMo: "A version of the classical traveling salesman problem with time dependent service times is proposed . The proposed model can handle several types of service time functions e.g. linear and quadratic functions. The proposed model can handle several types of service time functions e.g. linear and quadratic functions. Numerical results obtained by implementing different linear service time functions are presented."
PGM + ELMo + Cov: "This paper introduces a version of the classical traveling salesman problem with time dependent service times. The proposed model can minimize the total route duration which consists of the total service time. We apply several classes of subtour elimination constraints and measure their effect on the performance of our model . Numerical results obtained by implementing different test instances are presented."

Fig. 6.1: The input consists of only the abstract of a paper from the CSPubSum dataset. The highlights generated by each of the four models are presented. The input abstract and the author-written research highlights are taken from `https://www.sciencedirect.com/science/article/pii/S037722171500702X`

authors. Another observation is that the models with **ELMo** embeddings display better linguistic quality with respect to grammatical syntax probably due to the contextual nature of the embeddings. For example, while the first model (**PGM**) contains a grammatically incorrect sentence like 'We study minimize the ...' and the second model (**PGM + Cov**) generates the incorrect sentence 'Results are conducted ...', the ELMo-based models do not display such issues. However, none of the models capture the last line "Comparative computational results indicate that a flow based formulation is superior the other three" of the highlights penned by the authors because it does not appear in the abstract.

Figure 6.2 depicts a similar comparison among the outputs of the four models for a different paper. Again, we notice that without the coverage mechanism, words are incorrectly repeated, while the coverage mechanism reduces repetition significantly. Grammatical correctness of ELMo-based models is also more than that of other models. Figure 6.3 shows the comparison of predicted research highlights generated by the models for the same paper when the input (for training and test) is the combination of a research paper's abstract, introduction, and conclusion. Observe the same phenomenon of repetitive words in absence of the coverage mechanism: **PGM + ELMo** repeats the word 'graphical' several times, which is fixed when the coverage mechanism is added. However, a careful comparison of the outputs of **PGM + ELMo + Cov** across Figures 6.2 and 6.3 show that the highlights produced in the first case are more meaningful. For example, the statement 'We modeled the image as a superpixel based markov random field segmentation framework' (in **PGM + ELMo + Cov** model in 6.3) is not semantically correct. It seems that using a larger input from the body of the

Abstract: "Accurate and effective cervical smear image segmentation is required for automated cervical cell analysis systems. Thus, we proposed a novel superpixel-based Markov random field (MRF) segmentation framework to acquire the nucleus, cytoplasm and image background of cell images. We seek to classify color non-overlapping superpixel-patches on one image for image segmentation. This model describes the whole image as an undirected probabilistic graphical model and was developed using an automatic label-map mechanism for determining nuclear, cytoplasmic and background regions. A gap-search algorithm was designed to enhance the model efficiency. Data show that the algorithms of our framework provide better accuracy for both real-world and the public Herlev datasets. Furthermore, the proposed gap-search algorithm of this model is much more faster than pixel-based and superpixel-based algorithms."
Author-written research highlights: "We proposed a novel gap search markov random field mrf for accurate cervical smear image segmentation. This method could acquire three regions nuclei cytoplasm and background automatically by a label map mechanism. The gap search algorithm is faster than other three algorithms in the experiments. A copy of source codes will be released as an open source project for continuing studies."
PGM: "A novel model is developed to classify color non overlapping superpixel patches . Segmentation is used for the first time for image segmentation . Gap search algorithm for one image segmentation and superpixel based algorithms. Results are much more faster than pixel based and superpixel based model. Algorithms are more faster than pixel of superpixel based algorithms."
PGM + Cov: " We seek to classify color non overlapping superpixel patches on one image image . Work presents an label map mechanism to acquire the whole image as the undirected probabilistic graphical model . Gap search algorithm was designed to enhance the model efficiency. Algorithms much more faster than pixel based and superpixel superpixel ."
PGM + ELMo: "We propose a novel superpixel based markov random field mrf segmentation framework . The model describes the nucleus cytoplasm and image background of cell images. A gap search algorithm is designed to solve the model efficiency. The gap search algorithm is much more faster than pixel based and superpixel based algorithms."
PGM + ELMo + Cov: "A novel image segmentation method for automated cervical cell analysis is proposed. The whole image as an undirected probabilistic graphical model is presented . A gap search method is proposed to solve the nuclear cytoplasmic and background regions . The proposed algorithm is much more faster than pixel based and superpixel based algorithms."

Fig. 6.2: The input is only the abstract of a paper from the CSPubSum dataset. The highlights generated by each of the four models are presented. The input paper is at `https://www.sciencedirect.com/science/article/pii/S0010482516300154`

paper, rather than only the abstract, confuses the model, resulting in lower-quality output. Recollect that the superior performance of abstract-only models is quantitatively communicated by Table 5.1, and we now see a specific qualitative example of the performance difference. None of the models, irrespective of the input type, includes the information in the sentence "A copy of source codes will be released as an open source project for continuing studies." mentioned as the last line of the author-written highlights.

Thus, in all cases, we observe that the research highlights generated by the pointer-generator with ELMo embeddings and the coverage mechanism with the abstract as input are semantically more meaningful and contains fewer syntactic issues compared to the outputs in other cases. However, we also note that the research highlights produced by the models do not agree perfectly with the author-written highlights when the latter contain information not included in the inputs to the model.

7. Conclusion and future works. We analyzed four variants of deep neural models with two different input cases on CSPubSum dataset for the task of research highlight generation. The best performance was obtained by the pointer-generator model that utilized ELMo embeddings and coverage mechanism. Our future aim is to delve into the reasons behind the enhancement of model performance with pretrained embeddings. We also intend to integrate additional techniques to enhance the quality of the generated research highlights, both semantically and syntactically.

Author-written research highlights: "We proposed a novel gap search markov random field mrf for accurate cervical smear image segmentation. This method could acquire three regions nuclei cytoplasm and background automatically by a label map mechanism. The gap search algorithm is faster than other three algorithms in the experiments. A copy of source codes will be released as an open source project for continuing studies."
PGM: "We propose a novel superpixel based markov segmentation model . we combined the superpixels with an undirected probabilistic model. we combined the superpixels with an automatic label map mechanism. we combined the superpixels with an automatic label map mechanism. "
PGM + Cov: A novel superpixel based markov random field segmentation model to acquire the nucleus cytoplasm and image background of cell images we combined the superpixels with an undirected probabilistic graphical model and connecting edges . First the connecting edges of the graph reflect spatial information between neighboring superpixels by local probabilistic. second the connecting edges of the graph reflect spatial information between neighboring superpixels by local ."
PGM + ELMo: "A probabilistic graphical graphical graphical graphical graphical graphical graphical graphical graphical graphical graphical graphical graphical graphical an automatic superpixel based mrf segmentation model is proposed to acquire the requisite accuracy of the proposed gap of the proposed gap a superpixel based mrf segmentation model is proposed to enhance the model efficiency. The proposed framework is developed using an automatic label map mechanism for determining nuclear cytoplasmic and background regions."
PGM + ELMo + Cov: "We proposed a probabilistic graphical framework to acquire the nucleus cytoplasm and background in cervical smear images . We modeled the image as a superpixel based markov random field segmentation framework. A gap search algorithm was developed using an automatic label map mechanism for determining the model efficiency."

Fig. 6.3: The input consists of the concatenation of the abstract, introduction, and conclusion of a paper from the CSPubSum dataset. The highlights generated by each of the four models are presented. The input paper is at `https://www.sciencedirect.com/science/article/pii/S0010482516300154`

REFERENCES

[1] Richard Van Noorden. Global scientific output doubles every nine years. *Nature news blog*, 2014.

[2] Lutz Bornmann, Robin Haunschild, and Rüdiger Mutz. Growth rates of modern science: A latent piecewise growth curve approach to model publication numbers from established and new literature databases. *Humanities and Social Sciences Communications*, 8(1):1–15, 2021.

[3] Hans Peter Luhn. The automatic creation of literature abstracts. *IBM Journal of Research and Development*, 2(2):159–165, 1958.

[4] Julian Kupiec, Jan Pedersen, and Francine Chen. A trainable document summarizer. In *Proceedings of the 18th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 68–73, 1995.

[5] Ramesh Nallapati, Bowen Zhou, Cicero dos Santos, Caglar Gulcehre, and Bing Xiang. Abstractive text summarization using sequence-to-sequence RNNs and beyond. In *Proceedings of the 20th SIGNLL Conference on Computational Natural Language Learning*, 280–290, Berlin, Germany, 2016.

[6] Ce Zhou, Qian Li, Chen Li, Jun Yu, Yixin Liu, Guangjing Wang, Kai Zhang, Cheng Ji, Qiben Yan, Lifang He, et al. A comprehensive survey on pretrained foundation models: A history from BERT to ChatGPT. *arXiv:2302.09419*, 2023.

[7] Chin-Yew Lin. Rouge: A package for automatic evaluation of summaries. In *Text summarization branches out*, 74–81, 2004.

[8] Satanjeev Banerjee and Alon Lavie. METEOR: An automatic metric for MT evaluation with improved correlation with human judgments. In *Proceedings of the ACL Workshop on Intrinsic and Extrinsic Evaluation Measures for Machine Translation and/or Summarization*, 65–72, 2005.

[9] Tianyi Zhang, Varsha Kishore, Felix Wu, Kilian Q Weinberger, and Yoav Artzi. BERTScore: Evaluating text generation with BERT. In *8th International Conference on Learning Representations, (ICLR 2020)*, 1–43, 2020.

[10] Wei Zhao, Maxime Peyrard, Fei Liu, Yang Gao, Christian M. Meyer, and Steffen Eger. MoverScore: Text generation evaluating with contextualized embeddings and earth mover distance. In *EMNLP-IJCNLP*, 563–578. Association for Computational Linguistics, November 2019.

[11] Phyllis B Baxendale. Machine-made index for technical literature—an experiment. *IBM Journal of research and development*, 2(4):354–361, 1958.

[12] Harold P Edmundson. New methods in automatic extracting. *Journal of the ACM (JACM)*, 16(2):264–285, 1969.

[13] Kavita Ganesan, ChengXiang Zhai, and Jiawei Han. Opinosis: A graph based approach to abstractive summarization of highly redundant opinions. 2010.

[14] Ilya Sutskever, Oriol Vinyals, and Quoc V Le. Sequence to sequence learning with neural networks. *Proceedings of the 27th International Conference on Neural Information Processing Systems (NIPS 2014)*, 3104–3112, 2014.

[15] Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. Neural machine translation by jointly learning to align and translate. In *Proceedings of the 3rd International Conference on Learning Representations (ICLR 2015)*, 1–15, 2015.

[16] Sumit Chopra, Michael Auli, and Alexander M Rush. Abstractive sentence summarization with attentive recurrent neural networks. In *Proceedings of the 2016 conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 93–98, 2016.

[17] A. See, Peter J. Liu, and Christopher D. Manning. Get to the point: Summarization with pointer-generator networks. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*, vol.1, 1073–1083, 2017.

[18] Sebastian Gehrmann, Yuntian Deng, and Alexander M Rush. Bottom-up abstractive summarization. *preprint arXiv:1808.10792*, 2018.

[19] Peter J. Liu, Mohammad Saleh, Etienne Pot, Ben Goodrich, Ryan Sepassi, Lukasz Kaiser, and Noam Shazeer. Generating wikipedia by summarizing long sequences. In *International Conference on Learning Representations*, 2018.

[20] Dang Trung Anh and Nguyen Thi Thu Trang. Abstractive text summarization using pointer-generator networks with pre-trained word embedding. In *Proceedings of the 10th International Symposium on Information and Communication Technology*, 473–478, 2019.

[21] Nouf Ibrahim Altmami and Mohamed El Bachir Menai. Automatic summarization of scientific articles: A survey. *Journal of King Saud University-Computer and Information Sciences*, 2020.

[22] Danish Contractor, Yufan Guo, and Anna Korhonen. Using argumentative zones for extractive summarization of scientific articles. In *Proceedings of COLING 2012*, 663–678, December 2012.

[23] Kazutaka Kinugawa and Yoshimasa Tsuruoka. A hierarchical neural extractive summarizer for academic papers. In *New Frontiers in Artificial Intelligence: JSAI-isAI Workshops, JURISIN, SKL, AI-Biz, LENLS, AAA, SCIDOCA, kNeXI, Tsukuba, Tokyo, November 13-15, 2017, Revised Selected Papers 9*, 339–354. Springer, 2018.

[24] Ed Collins, Isabelle Augenstein, and Sebastian Riedel. A supervised approach to extractive summarisation of scientific papers. In *Proceedings of the 21st Conference on Computational Natural Language Learning (CoNLL 2017)*, pages 195–205, Vancouver, Canada, August 2017.

[25] Amanuel Alambo, Cori Lohstroh, Erik Madaus, Swati Padhee, Brandy Foster, Tanvi Banerjee, Krishnaprasad Thirunarayan, and Michael Raymer. Topic-centric unsupervised multi-document summarization of scientific and news articles. In *2020 IEEE International Conference on Big Data (Big Data)*, 591–596, 2020.

[26] Luca Cagliero and Moreno La Quatra. Extracting highlights of scientific articles: A supervised summarization approach. *Expert Systems with Applications*, 160:113659, 2020.

[27] Tohida Rehman, Debarshi Kumar Sanyal, Samiran Chattopadhyay, Plaban Kumar Bhowmick, and Partha Pratim Das. Automatic generation of research highlights from scientific. In *2nd Workshop on Extraction and Evaluation of Knowledge Entities from Scientific Documents (EEKE 2021), collocated with JCDL 2021*, 2021.

[28] Jingqing Zhang, Yao Zhao, Mohammad Saleh, and Peter Liu. PEGASUS: Pre-training with extracted gap-sentences for abstractive summarization. In *Procs. of the International Conference on Machine Learning*, 11328–11339. PMLR, 2020.

[29] Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J Liu. Exploring the limits of transfer learning with a unified text-to-text transformer. *preprint arXiv:1910.10683*, 2019.

[30] Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Ves Stoyanov, and Luke Zettlemoyer. Bart: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. *preprint arXiv:1910.13461*, 2019.

[31] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, 4171–4186, Minneapolis, Minnesota, June 2019. Association for Computational Linguistics.

[32] Zhaopeng Tu, Zhengdong Lu, Yang Liu, Xiaohua Liu, and Hang Li. Modeling coverage for neural machine translation. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 76–85, August 2016.

[33] Matthew E. Peters, Mark Neumann, Mohit Iyyer, Matt Gardner, Christopher Clark, Kenton Lee, and Luke Zettlemoyer. Deep contextualized word representations. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, vol. 1, 2227–2237, June 2018.

[34] Jeffrey Pennington, Richard Socher, and Christopher Manning. GloVe: Global vectors for word representation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1532–1543, October 2014.

[35] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. Distributed representations of words and phrases and their compositionality. *Advances in Neural Information Processing Systems*, 26, 2013.

[36] Jeffrey Pennington, Richard Socher, and Christopher D Manning. GloVe: Global vectors for word representation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1532–1543, 2014.

[37] Matt Kusner, Yu Sun, Nicholas Kolkin, and Kilian Weinberger. From word embeddings to document distances. In *International Conference on Machine Learning*, 957–966. PMLR, 2015.

Edited by: Katarzyna Wasielewska
Received: Apr 12, 2023
Accepted: Jun 8, 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 LaTeX 2$_\varepsilon$ 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.