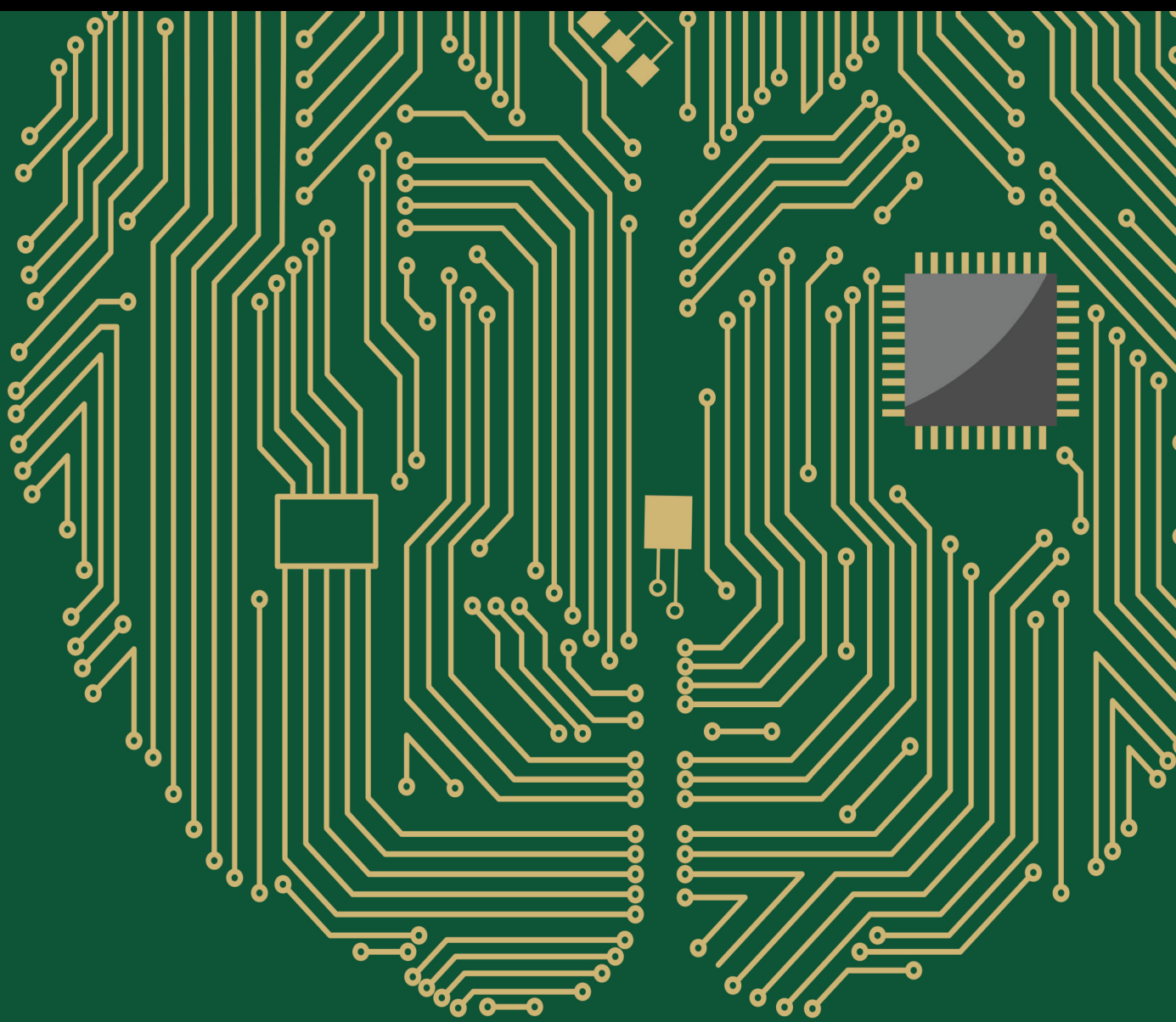


# Applications of Intelligent Systems on Health and the Environment

Lead Guest Editor: Habib Hamam

Guest Editors: Ahmed Cheriti, Ateeq Ur Rehman, and Inam Ullah





---

# **Applications of Intelligent Systems on Health and the Environment**

Computational Intelligence and Neuroscience

---

## **Applications of Intelligent Systems on Health and the Environment**

Lead Guest Editor: Habib Hamam

Guest Editors: Ahmed Cheriti, Ateeq Ur Rehman,  
and Inam Ullah



---

Copyright © 2023 Hindawi Limited. All rights reserved.

This is a special issue published in "Computational Intelligence and Neuroscience." All articles are open access articles distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

# Chief Editor

Andrzej Cichocki, Poland

## Associate Editors

Arnaud Delorme, France  
Cheng-Jian Lin , Taiwan  
Saeid Sanei, United Kingdom

## Academic Editors

Mohamed Abd Elaziz , Egypt  
Tariq Ahanger , Saudi Arabia  
Muhammad Ahmad, Pakistan  
Ricardo Aler , Spain  
Nouman Ali, Pakistan  
Pietro Aricò , Italy  
Lerina Aversano , Italy  
Ümit Ağbulut , Turkey  
Najib Ben Aoun , Saudi Arabia  
Surbhi Bhatia , Saudi Arabia  
Daniele Bibbo , Italy  
Vince D. Calhoun , USA  
Francesco Camastra, Italy  
Zhicheng Cao, China  
Hubert Cecotti , USA  
Jyotir Moy Chatterjee , Nepal  
Rupesh Chikara, USA  
Marta Cimitile, Italy  
Silvia Conforto , Italy  
Paolo Crippa , Italy  
Christian W. Dawson, United Kingdom  
Carmen De Maio , Italy  
Thomas DeMarse , USA  
Maria Jose Del Jesus, Spain  
Arnaud Delorme , France  
Anastasios D. Doulamis, Greece  
António Dourado , Portugal  
Sheng Du , China  
Said El Kafhali , Morocco  
Mohammad Reza Feizi Derakhshi , Iran  
Quanxi Feng, China  
Zhong-kai Feng, China  
Steven L. Fernandes, USA  
Agostino Forestiero , Italy  
Piotr Franaszczuk , USA  
Thippa Reddy Gadekallu , India  
Paolo Gastaldo , Italy  
Samanwoy Ghosh-Dastidar, USA

Manuel Graña , Spain  
Alberto Guillén , Spain  
Gaurav Gupta, India  
Rodolfo E. Haber , Spain  
Usman Habib , Pakistan  
Anandakumar Haldorai , India  
José Alfredo Hernández-Pérez , Mexico  
Luis Javier Herrera , Spain  
Alexander Hošovský , Slovakia  
Etienne Hugues, USA  
Nadeem Iqbal , Pakistan  
Sajad Jafari, Iran  
Abdul Rehman Javed , Pakistan  
Jing Jin , China  
Li Jin, United Kingdom  
Kanak Kalita, India  
Ryotaro Kamimura , Japan  
Pasi A. Karjalainen , Finland  
Anitha Karthikeyan, Saint Vincent and the Grenadines  
Elpida Keravnou , Cyprus  
Asif Irshad Khan , Saudi Arabia  
Muhammad Adnan Khan , Republic of Korea  
Abbas Khosravi, Australia  
Tai-hoon Kim, Republic of Korea  
Li-Wei Ko , Taiwan  
Raşit Köker , Turkey  
Deepika Koundal , India  
Sunil Kumar , India  
Fabio La Foresta, Italy  
Kuruva Lakshmana , India  
Maciej Lawrynczuk , Poland  
Jianli Liu , China  
Giosuè Lo Bosco , Italy  
Andrea Loddo , Italy  
Kezhi Mao, Singapore  
Paolo Massobrio , Italy  
Gerard McKee, Nigeria  
Mohit Mittal , France  
Paulo Moura Oliveira , Portugal  
Debajyoti Mukhopadhyay , India  
Xin Ning , China  
Nasimul Noman , Australia  
Fivos Panetsos , Spain

Evgeniya Pankratova , Russia  
Rocío Pérez de Prado , Spain  
Francesco Pistolesi , Italy  
Alessandro Sebastian Podda , Italy  
David M Powers, Australia  
Radu-Emil Precup, Romania  
Lorenzo Putzu, Italy  
S P Raja, India  
Dr.Anand Singh Rajawat , India  
Simone Ranaldi , Italy  
Upaka Rathnayake, Sri Lanka  
Navid Razmjooy, Iran  
Carlo Ricciardi, Italy  
Jatinderkumar R. Saini , India  
Sandhya Samarasinghe , New Zealand  
Friedhelm Schwenker, Germany  
Mijanur Rahaman Seikh, India  
Tapan Senapati , China  
Mohammed Shuaib , Malaysia  
Kamran Siddique , USA  
Gaurav Singal, India  
Akansha Singh , India  
Chiranjibi Sitaula , Australia  
Neelakandan Subramani, India  
Le Sun, China  
Rawia Tahrir , Iraq  
Binhua Tang , China  
Carlos M. Travieso-González , Spain  
Vinh Truong Hoang , Vietnam  
Fath U Min Ullah , Republic of Korea  
Pablo Varona , Spain  
Roberto A. Vazquez , Mexico  
Mario Versaci, Italy  
Gennaro Vessio , Italy  
Ivan Volosyak , Germany  
Leyi Wei , China  
Jianghui Wen, China  
Lingwei Xu , China  
Cornelio Yáñez-Márquez, Mexico  
Zaher Mundher Yaseen, Iraq  
Yugen Yi , China  
Qiangqiang Yuan , China  
Miaolei Zhou , China  
Michal Zochowski, USA  
Rodolfo Zunino, Italy

# Contents

## **Retracted: An AI-Driven Hybrid Framework for Intrusion Detection in IoT-Enabled E-Health**

Computational Intelligence and Neuroscience




Retraction (1 page), Article ID 9795813, Volume 2023 (2023)

## **Retracted: Artificial Bee Colony Based Gabor Parameters Optimizer (ABC-GPO) for Modulation Classification**

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9848575, Volume 2023 (2023)



## **Machine Learning Techniques for Antimicrobial Resistance Prediction of *Pseudomonas Aeruginosa* from Whole Genome Sequence Data**

Sohail M. Noman , Muhammad Zeeshan , Jehangir Arshad , Melkamu Deressa Amentie ,

Muhammad Shafiq , Yumeng Yuan , Mi Zeng , Xin Li , Qingdong Xie , and Xiaoyang Jiao 

Research Article (11 pages), Article ID 5236168, Volume 2023 (2023)


## **Supervised Machine Learning Based Noninvasive Prediction of Atrial Flutter Mechanism from P-to-P Interval Variability under Imbalanced Dataset Conditions**

Muhammad Usman Gul , Muhammad Haziq Kamarul Azman, Kushsairy Abdul Kadir , Jawad Ali

Shah, and Seada Hussien 

Research Article (18 pages), Article ID 8162325, Volume 2023 (2023)

## **A Privacy Protection Framework for Medical Image Security without Key Dependency Based on Visual Cryptography and Trusted Computing**

Denghui Zhang, Lijing Ren, Muhammad Shafiq , and Zhaoquan Gu

Research Article (11 pages), Article ID 6758406, Volume 2023 (2023)

## **An Explainable Statistical Method for Seizure Prediction Using Brain Functional Connectivity from EEG**

Hao Chen , Taoyun Ji , Xiang Zhan , Xiaoxin Liu , Guojing Yu , Wen Wang , Yuwu Jiang ,

and Xiao-Hua Zhou 

Research Article (8 pages), Article ID 2183562, Volume 2022 (2022)

## **The Diagnostic Value of Artificial Intelligence Ultrasound S-Detect Technology for Thyroid Nodules**


Peizhen Huang, Bin Zheng, Mengyi Li, Lin Xu, Sajjad Rabbani, Abdulillah Mohammad Mayet, Chengchun

Chen , Beishu Zhan , and He Jun 

Research Article (7 pages), Article ID 3656572, Volume 2022 (2022)






## **Development and Validation of Embedded Device for Electrocardiogram Arrhythmia Empowered with Transfer Learning**

Rizwana Naz Asif, Sagheer Abbas , Muhammad Adnan Khan , Atta-ur-Rahman , Kiran Sultan,


Maqsood Mahmud, and Amir Mosavi 

Research Article (15 pages), Article ID 5054641, Volume 2022 (2022)




**Understanding the User-Generated Geographic Information by Utilizing Big Data Analytics for Health Care**

Hidayat Ullah , Alaa Ali Hameed , Sanam Shahla Rizvi , Akhtar Jamil , and Se Jin Kwon   
Research Article (14 pages), Article ID 2532580, Volume 2022 (2022)





**Trends in Intelligent and AI-Based Software Engineering Processes: A Deep Learning-Based Software Process Model Recommendation Method**

Fahad H. Alshammari   
Research Article (11 pages), Article ID 1960684, Volume 2022 (2022)




**Botnet Attack Detection in IoT Using Machine Learning**

Khalid Alissa , Tahir Alyas , Kashif Zafar, Qaiser Abbas, Nadia Tabassum, and Shadman Sakib   
Research Article (14 pages), Article ID 4515642, Volume 2022 (2022)



**[Retracted] Artificial Bee Colony Based Gabor Parameters Optimizer (ABC-GPO) for Modulation Classification**

Saad AlJubayrin, Mubashar Sarfraz , Sajjad A. Ghauri , Muhammad Rizwan Amirzada , and  
Teweldebrhan Mezgebo Kebedew   
Research Article (9 pages), Article ID 9464633, Volume 2022 (2022)

**A Novel Framework for Fog-Assisted Smart Healthcare System with Workload Optimization**

Ahmed A. H. Abdellatif , Aman Singh , Abdulaziz Aldribi, Arturo Ortega-Mansilla, and Muhammad Ibrahim   
Research Article (12 pages), Article ID 4174805, Volume 2022 (2022)



**Application of Deep Learning in College Physical Education Design under Flipped Classroom**

Jun Huang  and Dian Yu   
Research Article (9 pages), Article ID 7368771, Volume 2022 (2022)





**A Natural Language Processing (NLP) Evaluation on COVID-19 Rumour Dataset Using Deep Learning Techniques**

Rubia Fatima , Naila Samad Shaikh, Adnan Riaz , Sadique Ahmad , Mohammed A. El-Affendi ,  
Khaled A. Z. Alyamani, Muhammad Nabeel, Javed Ali Khan , Affan Yasin , and Rana M. Amir Latif   
Research Article (17 pages), Article ID 6561622, Volume 2022 (2022)

**Analysis of Smart Lung Tumour Detector and Stage Classifier Using Deep Learning Techniques with Internet of Things**

Shubham Joshi , Shraddha Viraj Pandit, Piyush Kumar Shukla, Atiah H. Almalki, Nashwan Adnan Othman, Adnan Alharbi, and Musah Alhassan   
Research Article (11 pages), Article ID 4608145, Volume 2022 (2022)

**Performance Evaluation of Bio Concrete by Cluster and Regression Analysis for Environment Protection**


Ashish Shukla, Nakul Gupta, Kunwar Raghvendra Singh, Pawan Kumar Verma , Mohit Bajaj , Arfat Ahmad Khan , and Frie Ayalew   
Research Article (12 pages), Article ID 4411876, Volume 2022 (2022)



## Contents





---

### **Detection of Subarachnoid Hemorrhage in Computed Tomography Using Association Rules Mining**

Hathal Salamah Alwageed 



Research Article (10 pages), Article ID 1133819, Volume 2022 (2022)

### **[Retracted] An AI-Driven Hybrid Framework for Intrusion Detection in IoT-Enabled E-Health**

Fazal Wahab , Yuhai Zhao, Danish Javeed , Mosleh Hmoud Al-Adhaileh, Shahab Ahmad Almaaytah, Wasiat Khan, Muhammad Shahid Saeed , and Rajeev Kumar Shah 

Research Article (11 pages), Article ID 6096289, Volume 2022 (2022)

### **An Intelligent Cost-Efficient System to Prevent the Improper Posture Hazards in Offices Using Machine Learning Algorithms**

Jehangir Arshad , Hafiza Mahnoor Asim, Muhammad Adil Ashraf, Mujtaba Hussain Jaffery ,

Khurram Shabih Zaidi , and Melkamu Deressa Amentie 

Research Article (9 pages), Article ID 7957148, Volume 2022 (2022)

### **A Cumulants-Based Human Brain Decoding**

Raheel Zafar , Muhammad Javvad ur Rehman , Sheraz Alam , Muhammad Arslan Khan, Asad Hussain , Rana Fayyaz Ahmad , Faruque Reza , and Rifat Jahan 

Research Article (12 pages), Article ID 6474515, Volume 2022 (2022)

## *Retraction*

# **Retracted: An AI-Driven Hybrid Framework for Intrusion Detection in IoT-Enabled E-Health**

### **Computational Intelligence and Neuroscience**

Received 28 November 2023; Accepted 28 November 2023; Published 29 November 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### **References**

- [1] F. Wahab, Y. Zhao, D. Javeed et al., “An AI-Driven Hybrid Framework for Intrusion Detection in IoT-Enabled E-Health,” *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 6096289, 11 pages, 2022.

## Retraction

# Retracted: Artificial Bee Colony Based Gabor Parameters Optimizer (ABC-GPO) for Modulation Classification

### Computational Intelligence and Neuroscience

Received 8 August 2023; Accepted 8 August 2023; Published 9 August 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.





The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] S. AlJubayrin, M. Sarfraz, S. A. Ghauri, M. R. Amirzada, and T. Mezgebo Kebedew, "Artificial Bee Colony Based Gabor Parameters Optimizer (ABC-GPO) for Modulation Classification," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 9464633, 9 pages, 2022.

## Research Article

# Machine Learning Techniques for Antimicrobial Resistance Prediction of *Pseudomonas Aeruginosa* from Whole Genome Sequence Data

Sohail M. Noman <sup>1</sup>, Muhammad Zeeshan <sup>2</sup>, Jehangir Arshad <sup>3</sup>,  
Melkamu Deressa Amentie <sup>4</sup>, Muhammad Shafiq <sup>1</sup>, Yumeng Yuan <sup>1</sup>, Mi Zeng <sup>1</sup>,  
Xin Li <sup>1</sup>, Qingdong Xie <sup>1</sup> and Xiaoyang Jiao <sup>1</sup>

<sup>1</sup>Department of Cell Biology and Genetics, Shantou University Medical College, Shantou, Guangdong 515041, China

<sup>2</sup>Department of Medicine and Surgery, Al-Nafees Medical College and Hospital, Isra University, Islamabad 44000, Pakistan

<sup>3</sup>Department of Electrical and Computer Engineering, Comsats University Islamabad, Lahore Campus 44000, Lahore, Pakistan

<sup>4</sup>Department of Information Technology, Assosa University, Assosa 5220, Ethiopia

Correspondence should be addressed to Xiaoyang Jiao; xyjiao@stu.edu.cn

Received 19 August 2022; Revised 21 October 2022; Accepted 2 February 2023; Published 1 March 2023

Academic Editor: Inam Ullah

Copyright © 2023 Sohail M. Noman et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Aim.** Due to the growing availability of genomic datasets, machine learning models have shown impressive diagnostic potential in identifying emerging and reemerging pathogens. This study aims to use machine learning techniques to develop and compare a model for predicting bacterial resistance to a panel of 12 classes of antibiotics using whole genome sequence (WGS) data of *Pseudomonas aeruginosa*. **Method.** A machine learning technique called Random Forest (RF) and BioWeka was used for classification accuracy assessment and logistic regression (LR) for statistical analysis. **Results.** Our results show 44.66% of isolates were resistant to twelve antimicrobial agents and 55.33% were sensitive. The mean classification accuracy was obtained  $\geq 98\%$  for BioWeka and  $\geq 96$  for RF on these families of antimicrobials. Where ampicillin was 99.31% and 94.00%, amoxicillin was 99.02% and 95.21%, meropenem was 98.27% and 96.63%, cefepime was 99.73% and 98.34%, fosfomycin was 96.44% and 99.23%, ceftazidime was 98.63% and 94.31%, chloramphenicol was 98.71% and 96.00%, erythromycin was 95.76% and 97.63%, tetracycline was 99.27% and 98.25%, gentamycin was 98.00% and 97.30%, butirosin was 99.57% and 98.03%, and ciprofloxacin was 96.17% and 98.97% with 10-fold-cross validation. In addition, out of twelve, eight drugs have found no false-positive and false-negative bacterial strains. **Conclusion.** The ability to accurately detect antibiotic resistance could help clinicians make educated decisions about empiric therapy based on the local antibiotic resistance pattern. Moreover, infection prevention may have major consequences if such prescribing practices become widespread for human health.

## 1. Introduction

Antimicrobial resistance (AMR) is one of the leading public health concerns of the 21st century, which hinders the ability to effectively treat and prevent a wide variety of bacterial, viral, and fungal infections [1]. AMR occurs when microorganisms (bacteria, viruses, fungi, and parasites) evolve and lose their sensitivity to existing treatments, making infections more challenging to treat and raising the risk of

disease transmission, severe illness, and death [2]. The rapid global spread of multi- and pan-resistant bacteria, also known as “superbugs,” is particularly concerning because these bacteria cause infections that cannot be treated with current antimicrobial medicines like antibiotics [3]. At least 1.27 million people died from AMR-related cases in 2019, according to the CDC (<https://www.cdc.gov/drugresistance/biggest-threats.html>). Over 2.8 million people in the United States year contract AMR, and over 35,000 people die

directly [4]. The most common multidrug-resistant bacteria globally are *Escherichia coli*, *Enterococcus faecium*, *Streptococcus*, *Klebsiella*, and *Pseudomonas aeruginosa*, and they are responsible for an estimated 250,000 annual infections and deaths [5]. For instance, the WHO priority pathogen list calls for new antibacterials to treat infections caused by *Pseudomonas aeruginosa* and carbapenem-resistant bacteria (CRE) [6]. There are currently 32 antibiotics in clinical development that target WHO priority pathogens, but only six of them can be considered truly innovative [7].

Various researchers have talked about the resistance prediction of antimicrobials [8]. This lack of treatment options often requires broad-spectrum antibiotics, which may be less effective or safe. Resistance also affects empirical treatment, in which a clinician chooses an antibiotic for an infection without obtaining microbiological results. This can lead to an underestimation of the risk associated with specific infections and the use of inappropriate antibiotics. A meta-analysis found that patients with *Enterobacteriaceae* resistance are five times more likely to delay receiving an effective therapy than patients infected by a susceptible strain [9, 10]. This may reduce the long-term effectiveness of antibiotics, delay access to effective treatments, increase treatment failure with complications, and increase fatality rates. Infections caused by resistant Gram-positive and Gram-negative bacteria increase hospital stays, surgery needs, and mortality [11].

Another study by Yamani et al., calculated the health burden of antibiotic-resistant bacteria (ARB) in European Union/European Economic Area (EU/EEA) countries in disability-adjusted life-years [12]. Their models were populated with estimated incidence from the European Antimicrobial Resistance Surveillance Network (EARS-Net) and the European Centre for Disease Prevention and Control (ECDC) point prevalence surveys of healthcare-associated infections and antimicrobial use in European acute care hospitals [13, 14]. Systematic reviews of published literature showed attributable case fatality and length of stay for antibiotic-resistant infections [15, 16]. In 2014, 671,689 infections occurred in EU/EEA countries [13]. This ratio increased globally between 2015 and 2022 [5, 10, 12]. Different ARB contribute variably to the global burden, so prevention and control strategies should be tailored to each country's needs. All countries must implement effective AMR strategies to combat antibiotic overuse and misuse [17]. All systemic antibiotics globally require a doctor's prescription. Most prescriptions are written in primary care, not secondary or tertiary [6].

In 2018, 74% of all antibiotics prescribed by the National Health Service (NHS) in England were for general practitioners (GPs) patients [18]. GPs are the most frequent antibiotic prescribers, so they focus on primary care literature. Nurse practitioners and community pharmacists play a key role. In the last 10 years, nurses' roles have expanded to include prescribing in many countries and are on the policy agenda in many more [19]. Nurse prescribing was introduced to better utilize the skills and knowledge of health professionals, improve medication access, and reduce the workload of doctors. In China, the number of

nurses qualified to prescribe has steadily risen over the last 5 years, and 31,000 nurses now have the same prescribing ability as doctors [20]. Pharmacists in China can register as independent prescribers, often specializing in diabetes prescriptions. More pharmacists work in secondary care than primary. Lastly, dentists are considered antibiotic prescribers because they write fewer prescriptions than general practitioners. Further, most antibiotic prescriptions are for respiratory, urinary, skin, or tooth infections [21]. In addition, most antibiotics are given for acute respiratory tract infections (RTIs) [13]. Some RTIs, such as community-acquired bacterial pneumonia, are treatable with antibiotics, but most acute RTIs are viral and self-limiting.

*P. aeruginosa* has high baseline antibiotic resistance and can acquire new resistance mechanisms through chromosomal mutations or horizontal gene transfer (HGT), increasing the risk of ineffective antibiotic treatment [22]. Mutations can cause a failed therapeutic outcome during treatment, while resistance increases mortality, hospital stays, and costs. When microorganisms become resistant to antimicrobials, standard treatments are often ineffective. Disc diffusion and minimum inhibitory concentration (MIC) are the most common antimicrobial susceptibility tests [23]. Identification of resistance-specific markers by PCR or microarray hybridization is useful for epidemiological purposes and the validation of phenotypic results. As DNA sequencing throughput and costs increase, whole-genome sequencing (WGS) becomes a viable option for routine resistance profile surveillance and identifying emerging resistances [24]. Pathogenic *P. aeruginosa* alters genome sequences and protein expression to resist. Resistance disrupts biochemical pathways and protein channels [25]. Antibiotic resistance and susceptibility must be linked to specific resistance genes; all genes in an isolate are added to predict susceptibility [26]. ResFinder, CARD, and Resfams predict genotypes from phenotypes [27]. More and more often, computational tools like machine-learning algorithms are used to build models correlating genomic variations with phenotypes [28]. Both a stimulus and an outcome are present in every supervised learning example. The algorithm will succeed only if it learns a model that faithfully transforms any input into the desired output.

Considering the above, the fundamental objective of this study was to develop an accurate phenotype prediction model against antimicrobials. For this purpose, machine learning approaches called bio-Weka [29], and random forest (RF), and logistic regression (LR) [30–32] were used on the data mining platform called Weka (v3.9.2) (an open source java-based software) [33–35] for acquiring classification accuracy assumptions to accurately predict the phenotypes against a panel of twelve antimicrobial agents, including ampicillin, amoxicillin, meropenem, cefepime, fosfomycin, ceftazidime, chloramphenicol, erythromycin, tetracycline, gentamycin, butirosin, and ciprofloxacin from whole genome sequence data of *P. aeruginosa*. Significantly, this study can further enhance the antimicrobial predictions of various bacterial agents in clinical trials.

## 2. Methods

**2.1. Data Collection.** The WGS reads of *Pseudomonas aeruginosa* and binary resistance phenotypes of antimicrobial agents utilized in this study were obtained by accession numbers provided in various studies, consisting of different countries, including China and 65 others (developed and under development), and downloaded from the open access repository called GenBank at NCBI (<https://www.ncbi.nlm.nih.gov/genbank/>), which is the NIH genetic DNA sequences database. All the descriptive information about the raw data is present in the Supplementary file. The metadata consists of various attributes, including genome name, NCBI taxon id, genome status, associated strains, GenBank accession numbers, country name, number of contigs, genome lengths, isolation sources, resistance genes, twelve antibiotics, and many more.

**2.2. Model Framework and Parameters.** In this study, antimicrobial resistance of *P. aeruginosa* was predicted using a data mining assessment framework by machine learning algorithms, as shown in Figure 1. There were a total of six stages involved in reaching these conclusions, including the following: objective; data collection and preparation; machine learning techniques on a data mining platform; model building; evaluation and assessment; and implications. Initially, we collected the data and did some preliminary preprocessing to pick the right attributes. Afterward, this data was used for analysis and assessment. Secondly, Weka (v3.9.2), “a java-based machine learning and data mining platform,” was used to measure and evaluate classifications with the most recent bio-Weka and RF plugins. In addition, the results of machine learning classifiers were used in logistic regression (LR) to evaluate the resistance phenotype assessment to twelve different antibiotic drugs, namely, ampicillin, amoxicillin, meropenem, cefepime, fosfomycin, ceftazidime, chloramphenicol, erythromycin, tetracycline, gentamycin, butirosin, and ciprofloxacin.

Furthermore, the data was divided into two sets (training set and testing set) by a ratio of 60:40. Overfitting was prevented by using 10-fold cross-validation, and training data were used further as efficiently as possible to determine the optimal hyperparameter settings. The training model’s evaluation results were based on an average of the hyperparameter values that fared best in the 10-fold cross-validation procedure. Sensitivity, specificity, accuracy, and precision were used to assess the model performance of bio-Weka and RF by equations (1)–(4). The number of strains that turned out to be resistant was the true positive (TP), the number of strains that turned out to be sensitive was the true negative (TN), and the number of strains that turned out to be resistant when they should have been sensitive was the false positive (FP), and the number of strains that should have been sensitive when they should have been resistant was the false negative (FN) [36].

$$\text{Sensitivity} = \frac{TP}{(TP + FN)}, \quad (1)$$

$$\text{Specificity} = \frac{TN}{(TN + FP)}, \quad (2)$$

$$\text{Accuracy} = \frac{(TP + TN)}{(TP + FN + TN + FP)}, \quad (3)$$

$$\text{Precision} = \frac{TP}{(TP + FP)}. \quad (4)$$

**2.3. BioWeka and Random Forest Prediction of Phenotypes Resistance.** Weka’s datasets are used and stored in a unique file format known as attribute relation file format (ARFF). Due to the wide variety of file types used for biological data, it implements a format-conversion input layer that can transform common file types into the ARFF format. Weka filters any classes that can be applied to a dataset to alter it, and bio-Weka has filters for working with biological sequences. It enabled us to compare and match sequences with BLAST and other sequence alignment tools. In addition, alignment-based classification was performed using auto alignment score evaluation schemes.

A java-based machine learning algorithm called bio-Weka and RF was used to perform the predictive modeling. The DSK (k-mer counting software) [37, 38] was used to generate K-mer profiles (abundance profiles of all unique words of length  $k$  in each genome) from the assembled contigs, with  $k = 31$ . This is a common length for analyzing bacterial genomes [39]. In order to create the dataset, the 31-mer profiles of all strains were combined using the combine kmers tool in SEER [40]. The combined 31-mer counts were converted into presence/absence matrices to be used for model training and prediction. 10-fold cross-validation was used to select the best conjunctive and/or disjunctive model with a maximum of ten rules for binary classification analysis (using S/NS phenotypes based on the two different breakpoints for each drug) [41, 42], which involved testing the suggested broad range of values for the trade-off hyperparameter to determine the optimal rule scoring function (<https://aldr061.github.io/kover/doclearning.html>). In addition, classification (BW-mC) and regression (BW-R) models were constructed from log<sub>2</sub> (MIC) data in bio-Weka and RF for the purpose of comparing the performance of binary classifiers to MIC prediction [29, 43].

Furthermore, the RF method uses a majority voting strategy (MVS) to classify samples based on the results of an ensemble of decision tree (DT) [44]. In other words, the RF method relies on the class indicated by the vast majority of the DT. Having a diverse ensemble of trees is essential for boosting RF performance with respect to a single DT. One way to achieve it is by using bootstrapping with replacement to generate the training set for developing each DT’s unique

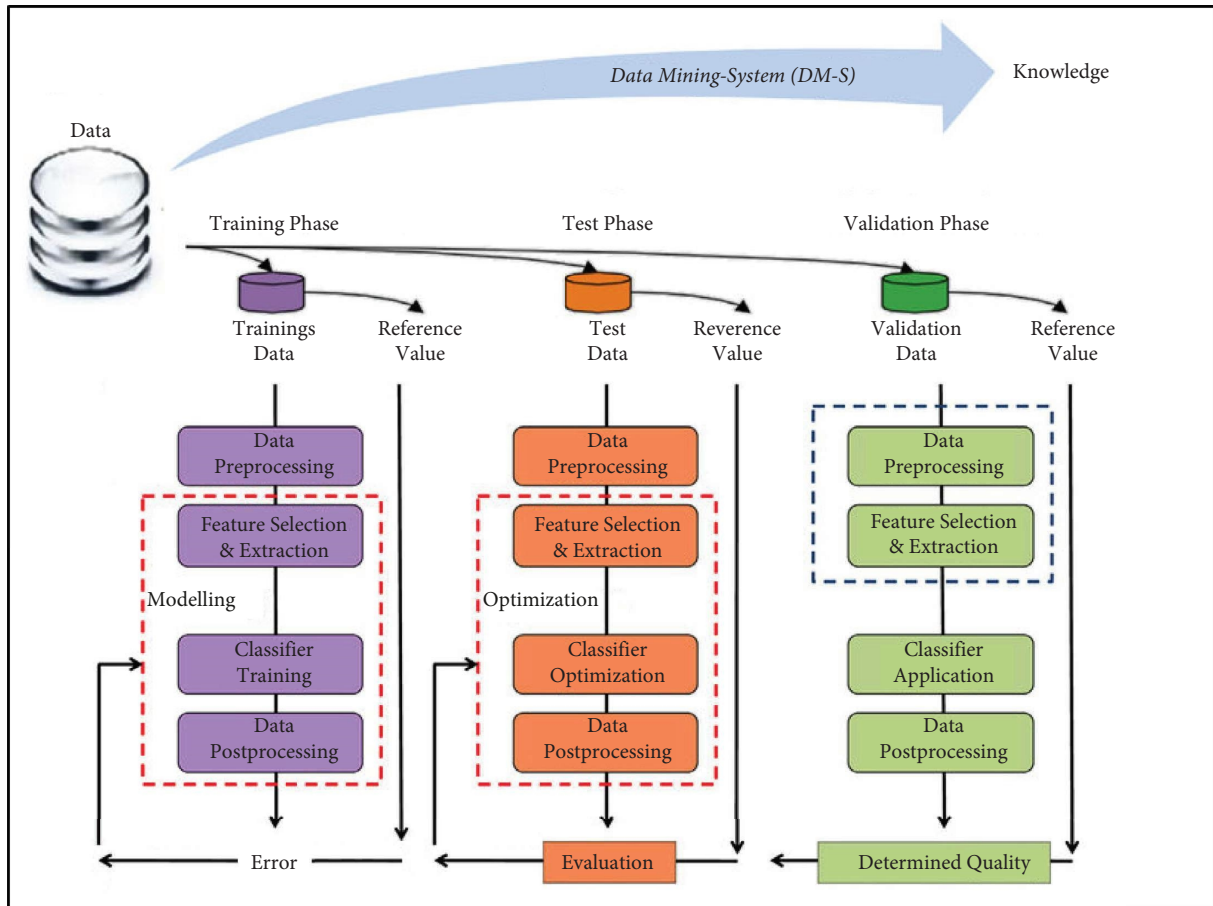


FIGURE 1: The data mining assessment framework used in this study.

feature set. However, features considered for splitting each node are not chosen from the full feature set but rather from a subset of features [45]. In addition, be aware that RF is more akin to an unintelligible black box model. In RF, as in individual DT, the CART algorithm is taken into account.

Multiple metrics were used to evaluate the model's efficacy, including sensitivity, specificity, accuracy, precision, and the overall bACC (the average of the sensitivity and specificity) [46]. Since the bACC represents false positive and false negative rates equally, regardless of the imbalance in the dataset, it was chosen as the overall measure of model performance. Two measures of MIC prediction accuracy were evaluated: firstly, the proportion of isolates for which the predicted MIC was identical to the phenotypic MIC (rounded to the nearest doubling dilution in the case of regression), and secondly, the proportion of isolates for which the predicted MIC was within one doubling dilution of the phenotypic MIC (1-tier accuracy). The MIC testing criteria for exact match rates and 1-tier accuracies have been removed to include predictions within 0.5 doubling dilutions or 1.5 doubling dilutions of the phenotypic MIC, respectively, to account for MIC variation [47]. Each analysis had 10 replicates, and the mean and 95% confidence intervals were calculated for all metrics. Mean bACC was

compared between replicate sets using two-tailed unpaired  $t$ -tests with logistic regression (LR) correction for unequal variance ( $\alpha = 0.05$ ) to assess differential model performance across datasets or methods. In addition,  $P$  values were calculated using the results of these unpaired  $t$ -tests.

**2.4. Regression Statistics.** Kappa statistics are reliable because they can be tested repeatedly [48, 49], ensuring that researchers have access to accurate, comprehensive data regarding research samples. It evaluates the predicted classification accuracy against a random classification [50]. We used a kappa statistic that relies on binary values, where 0 is considered as a null value and 1 represents the predicted outcome of the evaluation as in equation (5)–(7) [51]. It also serves as an indicator of the reliability of the evaluation. Not only that, but the LR variables help resolve the two-way binary classifications. When applied to the field of binary numbers, it makes predictions in the form of continuous values that allow for the preservation of sensitivity [36]. If the value is greater than the threshold (value > threshold), then the value assigned is 1; otherwise, the value measured is 0 as determined by the equations (8)–(11) [52].

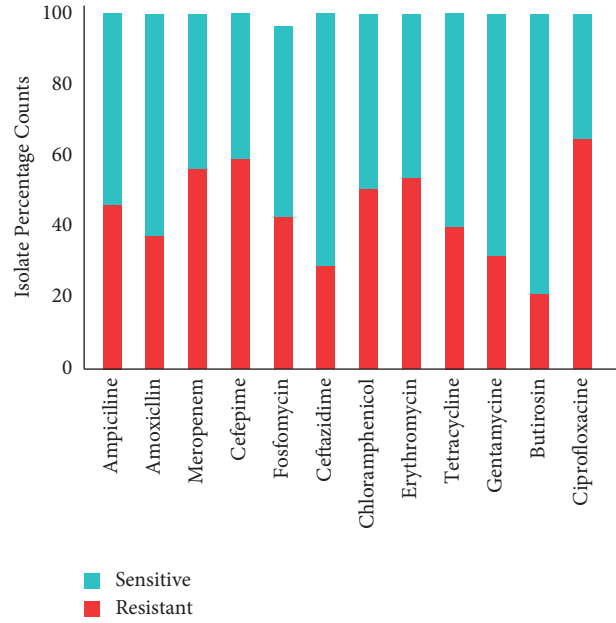


FIGURE 2: Number of resistant and sensitive isolate counts.

$$K = \frac{[P(A) - P(E)]}{[1 - P(E)]}, \quad (5)$$

$$P(A) = \left[ \frac{(TP + TN)}{N} \right], \quad (6)$$

$$P(E) = \left[ (TP + FN) * (TP + FP) * \frac{(TN + FN)}{N^2} \right], \quad (7)$$

$$P = \alpha + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_m X_m, \quad (8)$$

$$\sigma(x) \frac{1}{1 + e^{-x}} \in [0, 1], \quad (9)$$

$$\Pr(Y = +1|X) \sim \beta.X, \quad (10)$$

$$\begin{aligned} \Pr(Y = -1|X) \\ = 1 - \Pr(Y = +1|X). \end{aligned} \quad (11)$$

### 3. Results

A total of 1200 isolates of *P. aeruginosa* were included in this study, out of which 44.66% were resistant to 12 antimicrobial agents and 55.33% were sensitive, as shown in Figure 2. Of which 44.66% resistant isolates, 44 were resistant to ampicillin, 37 to amoxicillin, 58 to meropenem, 60 to cefepime, 45 to fosfomycin, 30 to ceftazidime, 52 to chloramphenicol, 58 to erythromycin, 39 to tetracycline, 30 to gentamycin, 20 to butirosin, and 63 to ciprofloxacin. In addition, of 55.33% of sensitive isolates, 56 were sensitive to ampicillin, 63 to amoxicillin, 42 to meropenem, 40 to cefepime, 55 to fosfomycin, 70 to ceftazidime, 48 to chloramphenicol, 42 to

erythromycin, 61 to tetracycline, 70 to gentamycin, 80 to butirosin, and 37 to ciprofloxacin, respectively. The most resistant genes to these twelve antimicrobial drugs were included *blaOXA-396*, *blaPAO*, *aph(3')-Iib*, *catB5*, *qacE*, *blaOXA-488*, *aac(6')-Ib-cr*, *aph(3')-Iia*, *aph(6)-Ic*, *aac(6')-Ib3*, *fosA*, *sul1*, *catB7*, *blaPAO*, *aac(3)-Ia*, *aac(6')-II*, *aph(3')-Iib*, *sul1catB7*, *blaPAO*, *blaOXA-396*, *blaOXA494*, *qacE*, *crpP*, *catB7*, *blaPAO*, and *blaOXA-488*. Furthermore, from the analysis total of 19,371,434, k-mers were obtained of length 31. Which were compared from the ResFinder k-mer genes database, and a range of (1,302,507) k-mers of *fosA*, *catB7*, *crpP*, *aac(6')-Ib-cr*, *fosA*, *tet(G)*, *aadA6*, *aph(3')-Iib*, *sul1*, *aph(3')-XV*, *aac(6')-Ib3*, *blaOXA-488*, *blaGES-13*, *blaGES-7*, *blaGES-5*, *blaGES-6*, *blaPAO*, *qacE*, *crpT*, *aph(3')-Iib*, *aadA13*, *blaOXA-50*, and *qacE* genes were detected in genome of 360 stains.

The accuracy percentage obtained from the results of BioWeka was more than 98% (as a mean percentage) including the training set and testing set, as shown in Figure 3 for all twelve antimicrobial drugs, namely, ampicillin, amoxicillin, meropenem, cefepime, fosfomycin, ceftazidime, chloramphenicol, erythromycin, tetracycline, gentamycin, butirosin, and ciprofloxacin with the confidence factor of 0.25% by 10-fold-cross validation. After the loop tests, the final mean accuracy for ampicillin was (99.31%), amoxicillin was (99.02%), meropenem was (98.27%), cefepime was (99.73%), fosfomycin was (96.44%), ceftazidime was (98.63%), chloramphenicol was (98.71%), erythromycin was (95.76%), tetracycline was (99.27%), gentamycin was (98.00%), butirosin was (99.57%), and ciprofloxacin was (96.17%).

In addition, Figure 4 shows the resulted classification accuracy percentage of RF algorithm in contrast to twelve antimicrobial drugs. The mean classification percentage was calculated more than 96% including the training set and



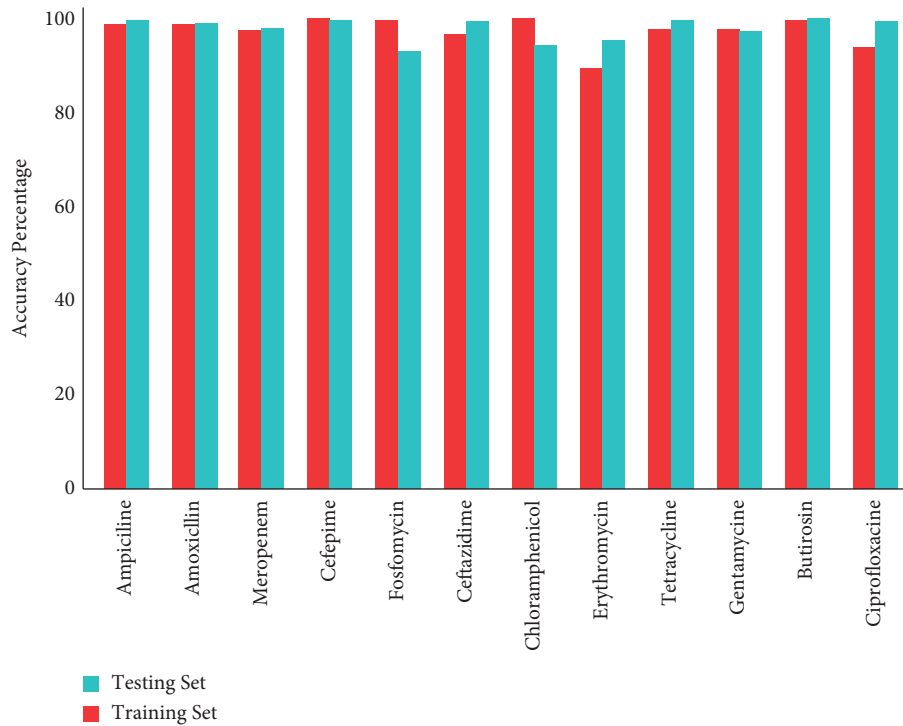


FIGURE 3: BioWeka classification accuracy percentage of the training set and testing set of twelve antimicrobial drugs.

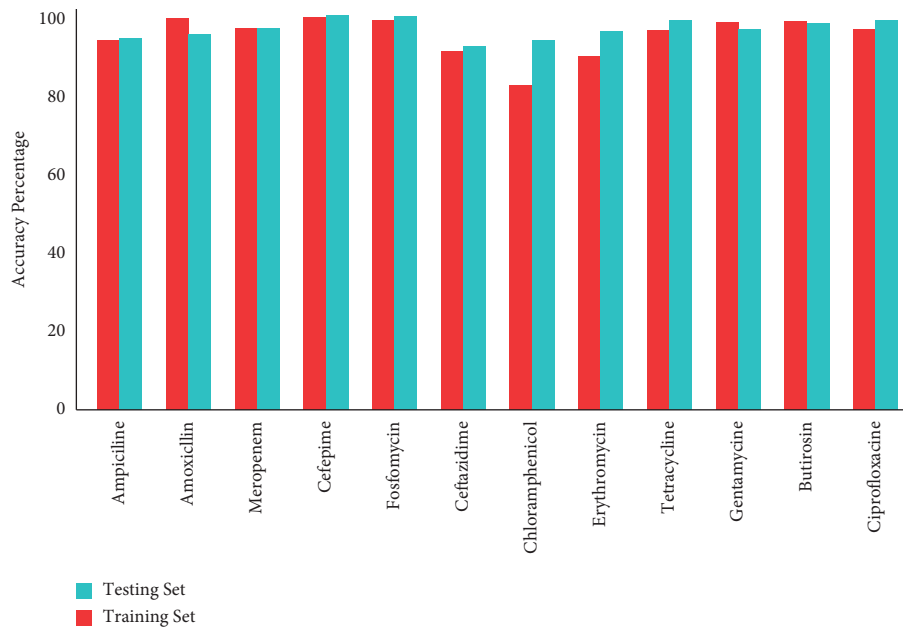


FIGURE 4: Random forest classification accuracy percentage of the training set and testing set of twelve antimicrobial drugs.

testing set, as shown in Figure 5. After the loop testing, the final accuracy by RF for ampicillin was (94.00%), amoxicillin was (95.21%), meropenem was (96.63%), cefepime was (98.34%), fosfomycin was (99.23%), ceftazidime was (94.31%), chloramphenicol was (96.00%), erythromycin was (97.63%), tetracycline was (98.25%), gentamycin was (97.30%), butirosin was (98.03%), and ciprofloxacin was

(98.97%). Furthermore, the standard deviation and average percentages of sensitivity, accuracy, precision, and specificity measured on the testing dataset are shown in Table 1. Our results of the testing dataset show that the antimicrobial drugs, namely ampicillin, amoxicillin, meropenem, cefepime, ceftazidime, tetracycline, butirosin, and ciprofloxacin, have no false-positive and false-negative bacterial strains.

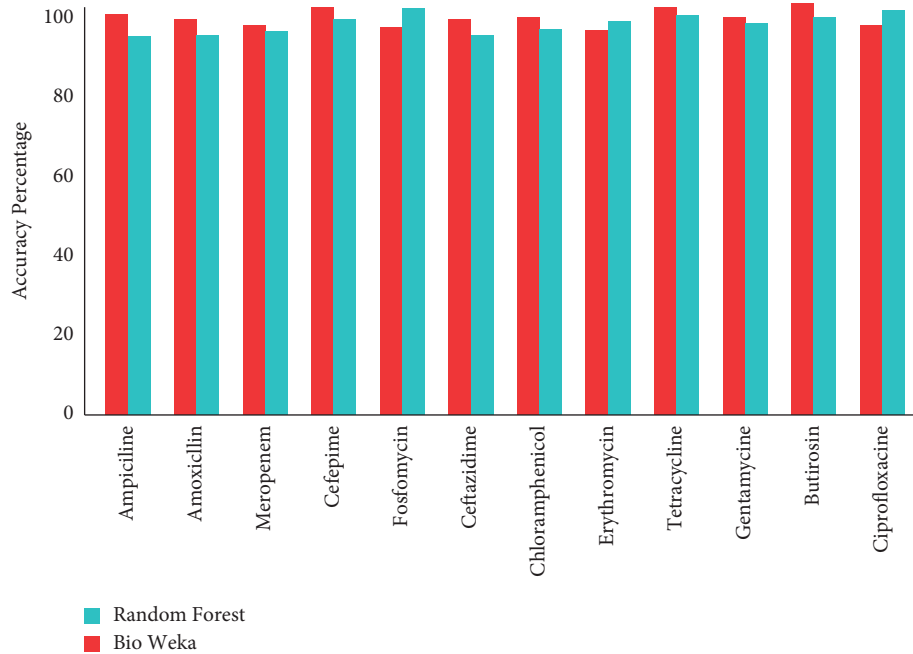


FIGURE 5: Mean accuracy percentage of random forest and BioWeka in comparison of twelve antimicrobial drugs.

#### 4. Discussion

A number of studies have highlighted the increasing global prevalence of antimicrobial resistance [12–16, 21, 24, 27, 53–57]. This is related to the challenges of treating bacterial infections, the consequences of which can be severe. *P. aeruginosa* is one of the most common bacterial species, and its families are responsible for some of the most dangerous infections ever seen in humans. There is a correlation between the resistance of these bacteria to multiple antibiotic classes and the severity of the infection, which complicates treatment. Antibiotic resistance among these microorganisms has been rising steadily over the years, and it is now common to find clinical samples resistant to multiple drugs. The development of antibiotic resistance causes doctors to delay administering the most effective treatment methods and prescribe a larger dosage of antibiotics than is necessary. This is particularly important in the intensive care unit, where patients' health conditions necessitate longer courses of antibiotics. The extensive use of expensive medical interventions, increased mortality rates, and lengthened hospital stays are all consequences of antimicrobial resistance [58]. Another topic of great interest is the need to prevent the spread of bacteria resistant to antibiotics and to identify them in advance so that patients can be isolated as soon as possible. Since this is the case, novel approaches must be proposed for detecting antimicrobial resistance and taking appropriate action without delay. In addition, gaining insight into the factors that contribute to the spread of nosocomial infections is possible by identifying relevant features.

In this paper, we propose a data mining strategy based on two machine learning techniques, namely, bio-Weka and RF with a statistical approach for detecting the antimicrobial

resistance of *P. aeruginosa* with different families of drugs. BioWeka and RF has shown that machine learning-based feature selection works with highly resulted accuracy as in Table 2. Consideration of antimicrobial drug resistance and susceptibility within data mining models and methods has been demonstrated to be useful in accelerating the workflow of clinical centers. Benefits for the individual, the healthcare system, and society may result from the early identification of patients at high risk of being resistant to one or more families of antibiotics. In addition, benefits include potential use in selecting the best antimicrobial treatment immediately.

Furthermore, the best performance achieved when testing this model strategy for resistance identification of antimicrobial drugs was a ROC area of 0.91 with a mean accuracy of more than 97% with all twelve drugs, indicating that our model can distinguish between the different classes of antibiotic susceptibility based solely on the type of the examined sample, the Gram stain classification of the pathogen, and prior antibiotic susceptibility testing results. We can foresee the sensitivity results from the various researchers using the model presented in this study. The ability to accurately detect antibiotic resistance could help clinicians make educated decisions about empiric therapy based on the local antibiotic resistance pattern. There may be major consequences for infection prevention if such prescribing practices become widespread.

The model proposed in this study has only the limitation with the process of filtering by 60:40 ratio with 10-fold cross-validation. If the ratios change then the accuracy and sensitivity of model might get affected. In addition, once the patient's clinical characteristics are added to the antimicrobial susceptibility dataset, the prediction performance of our model will significantly increase in terms of resistance

TABLE 1: Classification ratio of antimicrobial drugs against BioWeka and RF with phenotypes correlations.

Algorithm against drugs	Accuracy	Sensitivity	Specificity	Precision	F1 score	Kappa stats	Phenotype correlation	
Bio Weka classifications	Ampicillin	99.3 ± 0.0	86.0 ± 1.3	74.0 ± 2.3	1.0 ± 0.0	76.0 ± 3.2	91.0 ± 1.0 $p < 2.1e - 1$	
	Amoxicillin	99.0 ± 0.0	62.0 ± 1.2	88.3 ± 1.2	1.0 ± 0.0	77.0 ± 1.0	91.2 ± 1.0 $p < 2.1e - 1$	
	Meropenem	98.2 ± 0.0	88.0 ± 2.7	91.0 ± 1.0	1.0 ± 0.0	86.0 ± 2.5	89.3 ± 1.0 $p < 2.1e - 1$	
	Cefepime	99.7 ± 0.0	89.0 ± 1.0	89.0 ± 1.0	1.0 ± 0.0	77.0 ± 1.0	94.8 ± 1.0 $p < 2.1e - 1$	
	Fosfomycin	96.4 ± 0.0	77.0 ± 3.5	78.0 ± 2.1	1.0 ± 0.0	89.0 ± 1.0	97.6 ± 1.0 $p < 2.1e - 1$	
	Ceftazidime	98.6 ± 0.0	85.0 ± 14.2	86.0 ± 3.7	1.0 ± 0.0	88.6 ± 2.0	91.3 ± 1.0 $p < 2.1e - 1$	
	Chloramphenicol	98.7 ± 0.0	89.0 ± 2.1	78.0 ± 3.7	1.0 ± 0.0	91.9 ± 3.8	92.4 ± 1.0 $p < 2.1e - 1$	
	Erythromycin	95.7 ± 0.0	91.0 ± 12.3	86.0 ± 3.2	1.0 ± 0.0	87.0 ± 1.0	89.9 ± 1.0 $p < 2.1e - 1$	
	Tetracycline	99.2 ± 0.0	79.0 ± 1.7	89.0 ± 2.7	1.0 ± 0.0	79.0 ± 2.4	88.0 ± 1.0 $p < 2.1e - 1$	
	Gentamycin	98.0 ± 0.0	92.0 ± 2.5	77.0 ± 2.1	1.0 ± 0.0	81.0 ± 1.0	88.0 ± 1.0 $p < 2.1e - 1$	
	Butirosin	99.5 ± 0.0	88.0 ± 3.8	79.0 ± 12.1	1.0 ± 0.0	81.3 ± 2.7	87.6 ± 1.0 $p < 2.1e - 1$	
	Ciprofloxacin	96.1 ± 0.0	87.0 ± 2.4	91.0 ± 1.0	1.0 ± 0.0	85.0 ± 1.0	83.8 ± 1.0 $p < 2.1e - 1$	
	Random forest classification	Ampicillin	94.0 ± 0.0	81.5 ± 2.1	88.4 ± 1.0	1.0 ± 0.0	84.9 ± 1.0	81.1 ± 1.0 $p < 2.1e - 1$
		Amoxicillin	95.2 ± 0.0	88.4 ± 2.5	81.2 ± 2.1	1.0 ± 0.0	88.6 ± 1.0	84.3 ± 1.0 $p < 2.1e - 1$
Meropenem		96.6 ± 0.0	84.3 ± 3.6	73.9 ± 2.6	1.0 ± 0.0	87.1 ± 1.0	88.9 ± 1.0 $p < 2.1e - 1$	
Cefepime		98.3 ± 0.0	90.7 ± 2.2	77.0 ± 4.7	1.0 ± 0.0	82.5 ± 1.0	91.7 ± 1.0 $p < 2.1e - 1$	
Fosfomycin		99.2 ± 0.0	88.6 ± 2.3	76.8 ± 5.4	1.0 ± 0.0	77.7 ± 1.4	91.0 ± 1.0 $p < 2.1e - 1$	
Ceftazidime		94.3 ± 0.0	83.6 ± 2.1	83.7 ± 3.6	1.0 ± 0.0	79.0 ± 1.0	87.6 ± 1.0 $p < 2.1e - 1$	
Chloramphenicol		96.0 ± 0.0	89.7 ± 2.8	85.3 ± 2.9	1.0 ± 0.0	80.3 ± 2.7	84.9 ± 1.0 $p < 2.1e - 1$	
Erythromycin		97.6 ± 0.0	81.4 ± 4.6	82.6 ± 2.1	1.0 ± 0.0	78.7 ± 2.5	88.2 ± 1.0 $p < 2.1e - 1$	
Tetracycline		98.2 ± 0.0	83.9 ± 3.7	87.8 ± 3.1	1.0 ± 0.0	82.6 ± 1.0	91.9 ± 1.0 $p < 2.1e - 1$	
Gentamycin		97.3 ± 0.0	92.4 ± 2.6	79.6 ± 2.5	1.0 ± 0.0	89.4 ± 1.0	91.0 ± 1.0 $p < 2.1e - 1$	
Butirosin		98.0 ± 0.0	90.3 ± 3.1	81.9 ± 1.7	1.0 ± 0.0	86.3 ± 3.1	97.6 ± 1.0 $p < 2.1e - 1$	
Ciprofloxacin		98.9 ± 0.0	82.5 ± 3.5	88.6 ± 1.0	1.0 ± 0.0	81.2 ± 1.0	94.3 ± 1.0 $p < 2.1e - 1$	

TABLE 2: Our machine learning resulted model accuracy percentage comparison with recent studies.

Methods	Accuracy (%)	References
BioWeka	≥98	This paper
Random forest	≥96	
Support vector machine (SVM)	≥95	[59]
Set covering machine (SCM)	≥96	[59]
Logistic regression (LR)	≥93	[44]
Decision tree (DT)	≥95	[44]
Random forest (RF)	≥97	[44]
Multi-layer perceptron (MLP)	≥91	[44]

prediction accuracy to different drugs. However, still, any such inclusion must incur the cost of retrieving the relevant data, which may be an exercise that involves a number of healthcare units, thereby increasing communication costs and complicating the need to align protocols that may operate across departments. After incurring such information, it is important to evaluate how well the additional knowledge acquired in terms of the improved accuracy metrics of the model can be incorporated into the practice of the hospital physicians, who may need to reevaluate their decision-making processes in the context of supporting or contradicting recommendations from a decision support system. To sum up, we think of this study as a node on a spectrum of cost-effectiveness studies that data mining approaches and machine learning techniques will spark in the healthcare industry.

### Data Availability

All data used in this study can be found in the Supplementary file associated with this article, or it can also be made available upon request to the first author or corresponding author.

### Consent

Not applicable.

### Conflicts of Interest

The authors declare that they have no conflicts of interest.

### Authors' Contributions

Sohail M. Noman was responsible for conceptualization, methodology, empirical estimations, writing, and drafting of the original draft by. Supervision was performed by Xiaoyang Jiao. Sohail M. Noman, Muhammad Shafiq, Yumeng Yuan, Mi Zeng, Qingdong Xie, and Xin Li performed data collection. Sohail M. Noman, Muhammad Zeeshan, Jehangir Arshad, and Melkamu Deressa Amentie performed review and editing. All authors have read and approved the final manuscript.

### Acknowledgments

The authors would like to thank our colleagues who contributed in the study. This research work has been supported

by the Natural Science Foundation of China (NSFC) for young international scientists (Grant no. 42150410383); and the 2020 Li Ka Shing Foundation Cross-Disciplinary Research Grant (Project no. 2020LKSFG03E).

### Supplementary Materials

All descriptive information about the raw data is present in the Supplementary file. (*Supplementary Materials*)

### References

- [1] S. Druge, S. Ruiz, F. Vardon-Bouines et al., "Risk factors and the resistance mechanisms involved in *Pseudomonas aeruginosa* mutation in critically ill patients," *Journal of Intensive Care Medicine*, vol. 7, pp. 36–39, 2019.
- [2] S. Mohanty, B. Baliyarsingh, and S. Kumar Nayak, "Antimicrobial Resistance in *Pseudomonas aeruginosa*: A Concise Review," *Antimicrobial Resistance-A One Health Perspective*, IntechOpen, London, UK, pp. 1–11, 2021.
- [3] Z. Pang, R. Raudonis, B. R. Glick, T. J. Lin, and Z. Cheng, "Antibiotic resistance in *Pseudomonas aeruginosa*: mechanisms and alternative therapeutic strategies," *Biotechnology Advances*, vol. 37, no. 1, pp. 177–192, 2019.
- [4] R. F. Langendonk, D. R. Neill, and J. L. Fothergill, "The building blocks of antimicrobial resistance in *Pseudomonas aeruginosa*: implications for current resistance-breaking therapies," *Frontiers in Cellular and Infection Microbiology*, vol. 11, p. 307, 2021.
- [5] F. Prestinaci, P. Pezzotti, and A. Pantosti, "Antimicrobial resistance: a global multifaceted phenomenon," *Pathogens and Global Health*, vol. 109, no. 7, pp. 309–318, 2015.
- [6] E. Drenkard, "Antimicrobial resistance of *Pseudomonas aeruginosa* biofilms," *Microbes and Infection*, vol. 5, no. 13, pp. 1213–1219, 2003.
- [7] S. M. Noman, M. Shafiq, S. Bibi et al., "Exploring antibiotic resistance genes, mobile gene elements, and virulence gene factors in an urban freshwater samples using metagenomic analysis," *Environmental Science & Pollution Research*, vol. 30, no. 2, pp. 2977–2990, 2022.
- [8] M. C. Goodyear, N. E. Garnier, R. C. Levesque, and C. M. Khursigara, "Liverpool epidemic strain isolates of *Pseudomonas aeruginosa* display high levels of antimicrobial resistance during both planktonic and biofilm growth," *Microbiology Spectrum*, vol. 10, no. 3, e102514 pages, Article ID 1024222, 2022.
- [9] H. Mekonnen, A. Seid, G. Molla Fenta, and T. Gebrecherkos, "Antimicrobial resistance profiles and associated factors of *Acinetobacter* and *Pseudomonas aeruginosa* nosocomial infection among patients admitted at Dessie comprehensive

- specialized Hospital, North-East Ethiopia. A cross-sectional study,” *PLoS One*, vol. 16, no. 11, Article ID e0257272, 2021.
- [10] J. E. Moore, B. C. Millar, M. Ollman-Selinger, and L. Cambridge, “The role of suboptimal concentrations of nebulized tobramycin in driving antimicrobial resistance in *Pseudomonas aeruginosa* isolates in cystic fibrosis,” *Respiratory Care*, vol. 66, no. 9, pp. 1446–1457, 2021.
  - [11] M. Shafiq, S. U. Rahman, H. Bilal et al., “Incidence and molecular characterization of ESBL-producing and colistin-resistant *Escherichia coli* isolates recovered from healthy food-producing animals in Pakistan,” *Journal of Applied Microbiology*, vol. 133, no. 3, pp. 1169–1182, 2022.
  - [12] L. Yamani, A. Alamri, A. Alsultan, S. Alfifi, M. A. Ansari, and A. Alnimr, “Inverse correlation between biofilm production efficiency and antimicrobial resistance in clinical isolates of *Pseudomonas aeruginosa*,” *Microbial Pathogenesis*, vol. 157, Article ID 104989, 2021.
  - [13] A. Thacharodi and I. L. Lamont, “Aminoglycoside resistance in *Pseudomonas aeruginosa*: the contribution of the MexXY-OprM efflux pump varies between isolates,” *Journal of Medical Microbiology*, vol. 71, no. 6, p. 1563, 2022.
  - [14] S. Araújo Ma dos, J. S. Rodrigues, T. de L. G. F. Lobo, and F. C. de A. Maranhão, “Healthcare-associated infections by *Pseudomonas aeruginosa* and antimicrobial resistance in a public hospital from alagoas (Brazil),” *Jornal Brasileiro de Patologia e Medicina Laboratorial*, vol. 58, pp. 1–11, 2022.
  - [15] J. P. Lynch, G. G. Zhanel, and G. G. Zhanel, “*Pseudomonas aeruginosa* pneumonia: evolution of antimicrobial resistance and implications for therapy,” *Seminars in Respiratory and Critical Care Medicine*, vol. 43, no. 02, pp. 191–218, 2022.
  - [16] D. E. Madden, O. Olagoke, T. Baird et al., “Express yourself: quantitative real-time PCR assays for rapid chromosomal antimicrobial resistance detection in *Pseudomonas aeruginosa*,” *Antimicrobial Agents and Chemotherapy*, vol. 66, no. 5, 9 pages, Article ID e0020422, 2022.
  - [17] Y. Yuan, Y. Chen, F. Yao et al., “Microbiomes and resistomes in biopsy tissue and intestinal lavage fluid of colorectal cancer,” *Frontiers in Cell and Developmental Biology*, vol. 9, Article ID 736994, 2021.
  - [18] P. D. Tamma, S. L. Aitken, R. A. Bonomo, A. J. Mathers, D. van Duin, and C. J. Clancy, “Infectious diseases society of America guidance on the treatment of extended-spectrum  $\beta$ -lactamase producing enterobacterales (ESBL-E), carbapenem-resistant enterobacterales (CRE), and *Pseudomonas aeruginosa* with difficult-to-treat resistance (DTR-P. aeruginosa),” *Clinical Infectious Diseases*, vol. 72, no. 7, pp. e169–e183, 2021.
  - [19] M. Gajdác, Z. Baráth, K. Kárpáti et al., “No correlation between biofilm formation, virulence factors, and antibiotic resistance in *Pseudomonas aeruginosa*: results from a laboratory-based in vitro study,” *Antibiotics*, vol. 10, no. 9, p. 1134, 2021.
  - [20] Z. Y. Wu, X. S. Wu, W. Y. Yao, X. F. Wang, Z. W. Quan, and W. Gong, “Pathogens’ distribution and changes of antimicrobial resistance in the bile of acute biliary tract infection patients,” *Zhonghua wai ke za zhi [Chinese journal of surgery]*, vol. 59, no. 1, pp. 24–31, 2021.
  - [21] G. Boschetti, D. Sgarabotto, M. Meloni et al., “Antimicrobial resistance patterns in diabetic foot infections, an epidemiological study in northeastern Italy,” *Antibiotics*, vol. 10, p. 1241, 2021.
  - [22] A. Zahedi bialvaei, M. Rahbar, R. Hamidi-Farahani et al., “Expression of RND efflux pumps mediated antibiotic resistance in *Pseudomonas aeruginosa* clinical strains,” *Microbial Pathogenesis*, vol. 153, Article ID 104789, 2021.
  - [23] J. A. Karlowsky, A. J. Walkty, M. R. Baxter et al., “Vitro Activity of Cefiderocol against Extensively Drug-Resistant *Pseudomonas aeruginosa*: CANWARD, 2007 to 2019,” *Microbiology Spectrum*, vol. 10, no. 4, 2022.
  - [24] V. V. Mogasale, P. Saldanha, V. Pai, P. D. Rekha, and V. Mogasale, “A descriptive analysis of antimicrobial resistance patterns of WHO priority pathogens isolated in children from a tertiary care hospital in India,” *Scientific Reports*, vol. 11, p. 5116, 2021.
  - [25] C. Liu, M. Xu, X. Li, H. Dong, and L. Ming, “Trends in antimicrobial resistance in bloodstream infections at a large tertiary-care hospital in China: a 10-year retrospective study (2010–2019),” *Journal of Global Antimicrobial Resistance*, vol. 29, pp. 413–419, 2022.
  - [26] S. Vaz and M. Lall, “Potential public health impact of the development of antimicrobial resistance in clinical isolates of *Pseudomonas aeruginosa* on repeated exposure to biocides *In vitro*,” *Med J Dr DY Patil Vidyapeeth*, vol. 14, no. 1, p. 45, 2021.
  - [27] J. Soonthornsit, K. Pimwaraluck, N. Kongmuang, P. Pratya, and N. Phumthanakorn, “Molecular epidemiology of antimicrobial-resistant *Pseudomonas aeruginosa* in a veterinary teaching hospital environment,” *Veterinary Research Communications*, vol. 47, pp. 73–86, 2022.
  - [28] J. Qin, C. Zou, J. Tao et al., “Carbapenem resistant *Pseudomonas aeruginosa* infections in elderly patients: antimicrobial resistance profiles, risk factors and impact on clinical outcomes,” *Infection and Drug Resistance*, vol. 15, pp. 2301–2314, 2022.
  - [29] J. E. Gewehr, M. Szugat, and R. Zimmer, “BioWeka - extending the Weka framework for bioinformatics,” *Bioinformatics*, vol. 23, no. 5, pp. 651–653, 2007.
  - [30] M. N. Sohail, R. Jiadong, M. M. Uba et al., “Forecast Regression analysis for Diabetes Growth: an inclusive data mining approach,” *International Journal of Advanced Research in Computer Engineering*, vol. 7, pp. 715–721, 2018.
  - [31] S. M. Noman, J. Arshad, M. Zeeshan et al., “An empirical study on diabetes depression over distress evaluation using diagnosis statistical manual and chi-square method,” *International Journal of Environmental Research and Public Health*, vol. 18, no. 7, p. 3755, 2021.
  - [32] M. Uba Muhammad, R. Jiadong, M. N. Sohail, M. Irshad, M. Bilal, and A. A. Osi, “A logistic regression modeling on the prevalence of diabetes mellitus in the North Western Part of Nigeria,” *Benin Journal of Statistics-Uniben*, vol. 1, pp. 1–10, 2018.
  - [33] N. Sohail, J. Ren, I. Abir, M. Uba Muhammad, M. Bilal, and W. Iqbal, “WHY only data mining? A pilot study on inadequacy and domination of data mining technology why only data mining? A pilot study on inadequacy and domination of data mining technology (google scholar Library2019 view project WHY only data mining? A pilot st,” *International Journal of Recent Scientific Research*, vol. 9, Article ID 29066, 75 pages, 2018.
  - [34] M. N. Sohail, J. Ren, and M. Uba Muhammad, “A euclidean group assessment on semi-supervised clustering for health-care clinical implications based on real-life data,” *International Journal of Environmental Research and Public Health*, vol. 16, no. 9, pp. 1581–1593, 2019.
  - [35] M. M. Uba, R. Jiadong, M. N. Sohail, M. Irshad, and K. Yu, “Data mining process for predicting diabetes mellitus based model about other chronic diseases: a case study of the

- northwestern part of Nigeria,” *Healthcare Technology Letters*, vol. 6, no. 4, pp. 98–102, 2019.
- [36] M. U. Muhammad, R. Jiadong, N. S. Muhammad, and B. Nawaz, “Stratified diabetes mellitus prevalence for the Northwestern Nigerian States, a data mining approach,” *International Journal of Environmental Research and Public Health*, vol. 16, no. 21, p. 4089, 2019.
- [37] Y. Dufresne, T. Lemane, P. Marijon et al., “The K-mer File Format: a standardized and compact disk representation of sets of k-mers,” *Bioinformatics*, vol. 38, no. 18, pp. 4423–4425, 2022.
- [38] Y. Li, H. Patel, and Y. Lin, “Kmer2SNP: Reference-free Heterozygous SNP Calling Using K-Mer Frequency Distributions,” *Variant Calling: Methods and Protocols*, Springer, New York, NY, USA, pp. 257–265, 2022.
- [39] A. L. Hicks, N. Wheeler, L. Sánchez-Busó, J. L. Rakeman, S. R. Harris, and Y. H. Grad, “Evaluation of parameters affecting performance and reliability of machine learning-based antibiotic susceptibility testing from whole genome sequencing data,” *PLoS Computational Biology*, vol. 15, no. 9, 21 pages, Article ID e1007349, 2019.
- [40] X. Su, G. Jing, Y. Zhang, and S. Wu, “Method development for cross-study microbiome data mining: challenges and opportunities,” *Computational and Structural Biotechnology Journal*, vol. 18, pp. 2075–2080, 2020.
- [41] G. Feretzakis, E. Loupelis, A. Sakagianni et al., “Using machine learning techniques to aid empirical antibiotic therapy decisions in the intensive care unit of a general hospital in Greece,” *Antibiotics*, vol. 9, no. 2, p. 50, 2020.
- [42] H. Shakeri, V. Volkova, X. Wen et al., “Establishing statistical equivalence of data from different sampling approaches for assessment of bacterial phenotypic antimicrobial resistance,” *Applied and Environmental Microbiology*, vol. 84, no. 9, 17 pages, Article ID e, 2018.
- [43] G. Arango-Argoty, E. Garner, A. Pruden, L. S. Heath, P. Vikesland, and L. Zhang, “DeepARG: a deep learning approach for predicting antibiotic resistance genes from metagenomic data,” *Microbiome*, vol. 6, pp. 1–15, 2018.
- [44] S. Martínez-Agüero, I. Mora-Jiménez, J. Lérída-García, J. Álvarez-Rodríguez, and C. Soguero-Ruiz, “Machine learning techniques to identify antimicrobial resistance in the intensive care unit,” *Entropy*, vol. 21, no. 6, pp. 603–624, 2019.
- [45] W. Khan, L. Kong, S. M. Noman, B. Brekhna, and B. Brekhna, “A novel feature selection method via mining Markov blanket,” *Applied Intelligence*, vol. 1, pp. 1–24, 2022.
- [46] S. Suzuki, T. Horinouchi, and C. Furusawa, “Prediction of antibiotic resistance by gene expression profiles,” *Nature Communications*, vol. 5, no. 1, p. 5792, 2014.
- [47] G. Feretzakis, E. Loupelis, A. Sakagianni et al., “Using machine learning algorithms to predict antimicrobial resistance and assist empirical treatment,” *Studies in Health Technology and Informatics*, vol. 272, pp. 75–78, 2020.
- [48] M. U. Muhammad, R. Jiadong, N. S. Muhammad, M. Hussain, and I. Muhammad, “Principal component analysis of categorized polytomous variable-based classification of diabetes and other chronic diseases,” *International Journal of Environmental Research and Public Health*, vol. 16, no. 19, p. 3593, 2019.
- [49] M. N. Sohail, R. Jiadong, M. U. Muhammad, S. T. Chauhdary, J. Arshad, and A. J. Verghese, “An accurate clinical implication assessment for diabetes mellitus prevalence based on a study from Nigeria,” *Processes*, vol. 7, 2019.
- [50] M. N. Sohail, R. Jiadong, M. M. Uba et al., “A hybrid Forecast Cost Benefit Classification of diabetes mellitus prevalence based on epidemiological study on Real-life patient’s data,” *Scientific Reports*, vol. 9, Article ID 10103, 10110 pages, 2019.
- [51] M. U. Muhammad, O. E. Asiribo, and S. Muhammad, “Application of logistic regression modeling using fractional polynomials of grouped continuous covariates,” *Proceedings of the Nigeria Statistical Society*, vol. 1, pp. 144–147, 2017, [https://doi.org/http://nss.com.ng/2017\\_edited\\_proceedings](https://doi.org/http://nss.com.ng/2017_edited_proceedings).
- [52] M. N. Sohail, J. Ren, M. U. Muhammad et al., “Group covariates assessment on real-life diabetes patients by fractional polynomials: a study based on logistic regression modeling,” *Journal of Biotech Research*, vol. 10, pp. 116–125, 2019.
- [53] H. Bilal, M. N. Khan, T. Rehman, M. F. Hameed, and X. Yang, “Antibiotic resistance in Pakistan: a systematic review of past decade,” *BMC Infectious Diseases*, vol. 21, pp. 244–319, 2021.
- [54] M. Shafiq, J. Huang, J. M. Shah et al., “Characterization and virulence factors distribution of blaCTX-M and mcr-1 carrying *Escherichia coli* isolates from bovine mastitis,” *Journal of Applied Microbiology*, vol. 131, no. 2, pp. 634–646, 2021.
- [55] H. Bilal, F. Hameed, M. A. Khan, S. Khan, X. Yang, and T. U. Rehman, “Detection of mcr-1 gene in extended-spectrum  $\beta$ -lactamase-producing *Klebsiella pneumoniae* from human urine samples in Pakistan,” *Jundishapur Journal of Microbiology*, vol. 13, no. 4, pp. 13–21, 2020.
- [56] H. Bilal, T. U. Rehman, M. A. Khan et al., “Molecular epidemiology of mcr-1, blaKPC-2, and blaNDM-1 harboring clinically isolated *Escherichia coli* from Pakistan,” *Infection and Drug Resistance*, vol. 14, pp. 1467–1479, 2021.
- [57] H. Bilal, G. Zhang, T. Rehman et al., “First report of blaNDM-1 bearing IncX3 plasmid in clinically isolated ST11 *Klebsiella pneumoniae* from Pakistan,” *Microorganisms*, vol. 9, no. 5, p. 951, 2021.
- [58] M. Shafiq, J. Huang, S. Ur Rahman et al., “High incidence of multidrug-resistant *Escherichia coli* coharboring mcr-1 and blaCTX-M-15 recovered from pigs,” *Infection and Drug Resistance*, vol. 12, pp. 2135–2149, 2019.
- [59] Z. Liu, D. Deng, H. Lu et al., “Evaluation of machine learning models for predicting antimicrobial resistance of actinobacillus pleuropneumoniae from whole genome sequences,” *Frontiers in Microbiology*, vol. 11, pp. 48–57, 2020.

## Research Article

# Supervised Machine Learning Based Noninvasive Prediction of Atrial Flutter Mechanism from P-to-P Interval Variability under Imbalanced Dataset Conditions

Muhammad Usman Gul <sup>1</sup>, Muhammad Haziq Kamarul Azman,<sup>1</sup>  
Kushsairy Abdul Kadir <sup>1</sup>, Jawad Ali Shah,<sup>2</sup> and Seada Hussien <sup>3</sup>

<sup>1</sup>Universiti Kuala Lumpur, British Malaysian Institute, Kuala Lumpur, Malaysia

<sup>2</sup>Islamic International University, Islamabad, Pakistan

<sup>3</sup>School of Electrical and Computer Engineering, Haramaya Institute Technology, Diredawa 138, Ethiopia

Correspondence should be addressed to Kushsairy Abdul Kadir; [kushsairy@unikl.edu.my](mailto:kushsairy@unikl.edu.my) and Seada Hussien; [seada.hussen@aastu.edu.et](mailto:seada.hussen@aastu.edu.et)

Received 29 July 2022; Revised 1 October 2022; Accepted 12 October 2022; Published 1 March 2023

Academic Editor: Inam Ullah

Copyright © 2023 Muhammad Usman Gul et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Atrial flutter (AFL) is a common arrhythmia with two significant mechanisms, namely, focal (FAFL) and macroreentry (MAFL). Discrimination of the AFL mechanism through noninvasive techniques can improve radiofrequency ablation efficacy. This study aims to differentiate the AFL mechanism using a 12-lead surface electrocardiogram. P-P interval series variability is hypothesized to be different in FAFL and MAFL and may be useful for discrimination. 12-lead ECG signals were collected from 46 patients with known AFL mechanisms. Features for a proposed classifier are extracted through descriptive statistics of the interval series. On the other hand, the class ratio of MAFL and FAFL was 41 : 5, respectively, which was highly imbalanced. To resolve this, different data augmentation techniques (SMOTE, modified-SMOTE, and smoothed-bootstrap) have been applied on the interval series to generate synthetic interval series and minimize imbalance. Modification is introduced in the classic SMOTE technique (modified-SMOTE) to properly produce data samples from the original distribution. The characteristics of modified-SMOTE are found closer to the original dataset than the other two techniques based on the four validation criteria. The performance of the proposed model has been evaluated by three linear classifiers, namely, linear discriminant analysis (LDA), logistic regression (LOG), and support vector machine (SVM). Filter and wrapper methods have been used for selecting relevant features. The best average performance was achieved at 400% augmentation of the FAFL interval series (90.24% sensitivity, 49.50% specificity, and 76.88% accuracy) in the LOG classifier. The variation of consecutive P-wave intervals has been shown as an effective concept that differentiates FAFL from MAFL through the 12-lead surface ECG.

## 1. Introduction

Atrial flutter (AFL) is a common type of supraventricular tachycardia (SVT). Based on population studies, it is estimated that there will be annually 200,000 new AFL cases in the US alone [1]. AFL arrhythmia is characterized by electrical signals that regularly propagate along various conduction pathways within the myocardial tissue with self-sustaining mechanisms [2]. Recurrent sustained AFL can lead to significant symptoms

such as palpitations, fatigue, syncope, stroke, and even heart attack.

Atrial flutter can be classified into two different mechanisms, that is, focal AFL (FAFL) and macroreentry AFL (MAFL) according to the characteristics of the conduction propagation [3]. FAFL starts from a single spot, while MAFL is a process that circles a significant obstacle (see Figure 1 for an illustration).

An effective invasive treatment for AFL is radiofrequency catheter ablation (RFCA). It aims to create

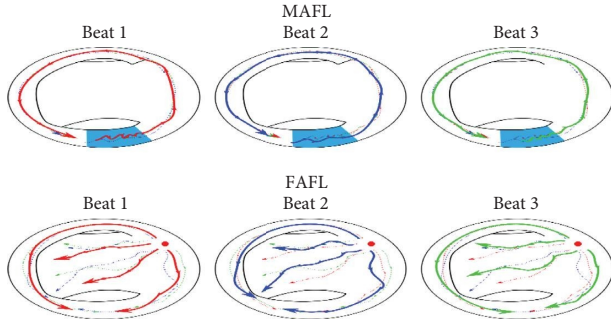


FIGURE 1: Hypothetical propagation of depolarization wavefront inside the atrium (depicted here: right atrium). The beat number indicates a sequence of atrial beats. Note the path similarity in MAFL, and the lack of it in FAFL.

a conduction barrier to block the reentry of the loop or to destroy ectopic pacemakers. In addition, the list of abbreviations is given in Table 1. During the procedure, intracardiac electrocardiograms, pacing maneuvers, and isochrone mapping are used to characterize the AFL mechanism during RFCA treatment [2, 4, 5]. This can only be performed once the catheters are introduced into the heart.

Although RFCA is preferred for AFL treatment due to its efficacy, its significant dependence on the electrophysiology study makes it a time-consuming and laborious treatment. Therefore, its efficacy can be improved before going under any invasive procedure if the characteristics of AFL (e.g., its mechanism) can be identified from some noninvasive techniques. The noninvasive 12-lead surface ECG is mostly used in clinical detection to differentiate the AFL from atrial fibrillation or normal ECG [6]. The noninvasive differentiation of the AFL mechanisms through 12-lead ECG would help clarify the ablation strategy and reduce the required time and resources in invasive cardiac mapping. However, it has been pointed out that differentiating AFL mechanisms from surface ECG can be a difficult task [7].

Previous attempts are highly dependent on delineation [8] and morphology [3, 9, 10] of atrial waves to discriminate the AFL mechanism through 12-lead surface ECG. Recently, the recurrence quantification analysis (RQA)-based model has been used to distinguish AFL mechanisms from a synthetically generated ECG dataset based on a computational model [11, 12]. Moreover, the authors highlighted the differences in ECG features of FAFL and MAFL.

Classifiers are tailored to differentiate two or many classes using advanced data processing techniques. Their hyperparameters are typically tuned by minimizing some cost function based on a set of given data. However, it most frequently disregards the issue of class prevalence. When the classes are imbalanced (one class more prevalent than others), the classifier’s decisions favor the majority class [13, 14]. This affects the classifier’s performance, and hence the resulting diagnosis. In medical science, datasets are often imbalanced due to, e.g., rarity of the disease, difficulty in obtaining data, or time and money constraints.

TABLE 1: List of abbreviations.

AFL	Atrial Flutter
CDF	Cumulative distribution function
CL	Cycle length
FAFL	Focal atrial flutter
GLRT	Generalized likelihood ratio test
MAFL	Macroreentrant atrial flutter
LDA	Linear discriminant analysis
LOG	Logistic regression
PWM	P-wave morphology
RFCA	Radiofrequency catheter ablation
RQA	Recurrence quantification analysis
SVM	Support vector machine
SVT	Supraventricular tachycardia

In this study, we propose a classifier model for discriminating the AFL mechanism using peak-to-peak successive atrial activities of 12-lead surface ECG through linear classifiers. Furthermore, the feature extraction of the proposed model is independent of the delineation and morphology of atrial activities.

Due to the class imbalance present in our working dataset, several data augmentation techniques have been used to counteract this. The synthetic minority oversampling technique (SMOTE) [15] is used widely in different fields that create synthetic samples from minority classes to balance the dataset and minimize bias, e.g., article [16]. However, SMOTE has been shown to shrink the variance of the original dataset and introduce correlation between samples [17]. In this study, we also propose a theoretical modification to the original algorithm in order to correct the shrunken variance issue.

The contribution made by the authors in this article is highlighted as follows: (i) methodological improvements to correct the variance shrinkage in classical SMOTE, (ii) a method to parametrize the generation of synthetic P-P intervals in terms of real data, (iii) a novel classifier to differentiate focal from macroreentrant AFL, and (iv) selection of several relevant features that discriminate FAFL from MAFL.

The organization of this article is outlined here. Section 2 describes the preprocessing of raw data after collection from the hospital and further elaborates on the hypothesis idea. Different standard data augmentation techniques and data generation protocols are explained as feature extraction and selection methodologies. Section 3 reports the results of data augmentation technique validation to identify the best augmentation rate. Section 4 discusses the results obtained in the section on the classification results before and after augmentation and the selection of relevant features. Section 5 concludes the article.

## 2. Materials and Methods

**2.1. Data Acquisition and Preprocessing.** The 61 patients who took part in this study were all under consideration for AFL ablation between January and December 2017 and were registered in a French hospital in Monaco, known as the Centre Hospitalier Princesse Grace. The patients’ demography with collection parameters is summarized in Table 2.



TABLE 2: Patients' demography.

Variable	Focal AFL	Macroreentrant AFL
Cycle length (msec)	230 ± 22.36	248.10 ± 39.38
Before cardiac surgery or ablation	3/5	22/41*
Left circuit	2/5	19/41
NonCTI	5/5	19/41**

(\*) three ECGs are undefined and (\*\*) one ECG is defined.

All ECGs were obtained during the ablation process using the acquisition system (Boston Scientific, USA). The ECGs were obtained using nine electrodes that were placed on the surface of the body. In contrast, the sampling frequency of the system was set to 2 kHz. The quantization level, which served as the analog to digital conversion resolution, was set to 16 bits. A supine angle was kept the same for all the patients on the surgical table.

In electrophysiological investigations conducted during the ablation procedure, mapping maneuvers establish the target for the ablated source, which distinguishes between the AFL mechanism and other mechanisms, enabling the accurate determination of the ablation spot. The report is made after the ablation. It provides information about the patient's condition, the AFL, the circuit, the orientation, and the direction of the circuit, among other things, which have been used as the ground truth in this study.

Consecutive P-waves are required in this study. The ratio between atria and ventricle conduction must be sufficiently large in order to identify these waves, hence we do not include ECGs whose conduction ratio is less than or equal to 2:1. Therefore, a limitation ratio is required between atria and ventricles to avoid nonconsecutive issues. With all these limitations, 5 FAFL and 41 MAFL valid records were obtained. These ECGs were then bandpass filtered (passband = [3:40] Hz) using two-stage high-pass and low-pass filters with a Chebyshev type II structure. A notch filter at 50 Hz is also used to remove powerline interference. The overall methodology of the proposed research structure is arranged in Figure 2, which can be clearly understood as the novel method used for discrimination of the AFL mechanism.

## 2.2. Calculating the Intervals of Consecutive P-Waves.

Atrial flutter can be characterized by electrical signals that repeatedly propagate along various physiological pathways [2, 18]. Macroreentrant and focal AFL have very different activation patterns in the endocardium. One is in the form of a stable reentrant loop, and the other is a point source from which depolarization originates and propagates throughout the entire atrial structure. As shown in Figure 1, the depolarization wavefront patterns in the two mechanisms are quite different. On one hand, it is expected that in MAFL, the stable circuit produces a stable ECG pattern without much variation. On the other hand, centrifugal depolarization in FAFL cannot guarantee that similar paths will be encountered by the wavefront. This instability will translate into varying ECG patterns.

It is hypothesized that the mechanism of AFL can be differentiated from 12-lead surface ECG based on the cycle length variability of the visible consecutive atrial activities (i.e., two or more than two P-waves within R-R interval). Notably, the intervals between P-wave peaks are hypothesized to be more variable in FAFL than in MAFL.

For each record, the lead containing the largest R-wave energy is selected, and the peaks of atrial activities (P-waves) have been identified by the GLRT method [19]. As shown in Figure 3, intervals between each P-wave have been measured and collected as a series. P-waves overlapped with T-waves have also been considered in this study and are estimated using the least square polynomial estimation [20]. Next, intervals above 300 ms are removed from the series, to ensure that false intervals not related to consecutive P-waves are removed. In total, 444 and 2546 intervals were obtained from 5 focal and 41 macroreentrant AFL ECGs, respectively.

**2.3. Synthetic Data Augmentation for Balancing.** The current dataset presents a heavily imbalanced class ratio of about 8:1 for MAFL against FAFL. One can expect the results to be biased towards MAFL (the majority class). Various kinds of data augmentation techniques have been proposed to overcome the imbalance issues present [21]. This study proposes a comparative study among three such techniques to conclude on which one is the best for the considered scenario. These techniques are SMOTE [15, 22], modified-SMOTE, and smoothed-bootstrap. The input to be augmented is the series of P-P intervals.

**2.3.1. SMOTE.** The main idea of SMOTE is to generate new synthetic interval data based on the linear combination of two interval data  $X_{j^{\text{th}}\text{Interval}}$  and  $X_{j^{\text{th}}\text{Interval}}^k$  where the latter is one of the  $k$ -nearest neighbors of the former. The synthetic data is then an interpolation within the sample space in a defined neighbourhood. The new synthetic interval is defined as

$$S_{j^{\text{th}}\text{Interval}} = X_{j^{\text{th}}\text{Interval}} + \alpha(X_{j^{\text{th}}\text{Interval}}^k - X_{j^{\text{th}}\text{Interval}}), \quad (1)$$

where  $S_{j^{\text{th}}\text{Interval}}$  is the synthetic interval and  $\alpha$  is a random number belonging to [0, 1]. The result is a synthetic interval data that is randomly generated along the line between  $X_{j^{\text{th}}\text{Interval}}$  and  $X_{j^{\text{th}}\text{Interval}}^k$ . In our scenario, we consider the neighborhood to encompass the totality of the dataset instead of a local neighborhood (i.e.,  $k = N - 1$  assuming there are  $N$  examples in the dataset).

**2.3.2. Modified-SMOTE.** Some theoretical properties of SMOTE for high-dimensional in-class imbalanced data have been discussed in [17]. One such property is that the synthetic samples have the same mean as the original dataset, but its variance is shrunk by a factor of 2/3. To counteract the shrunken variance, we propose the following modification to (1):

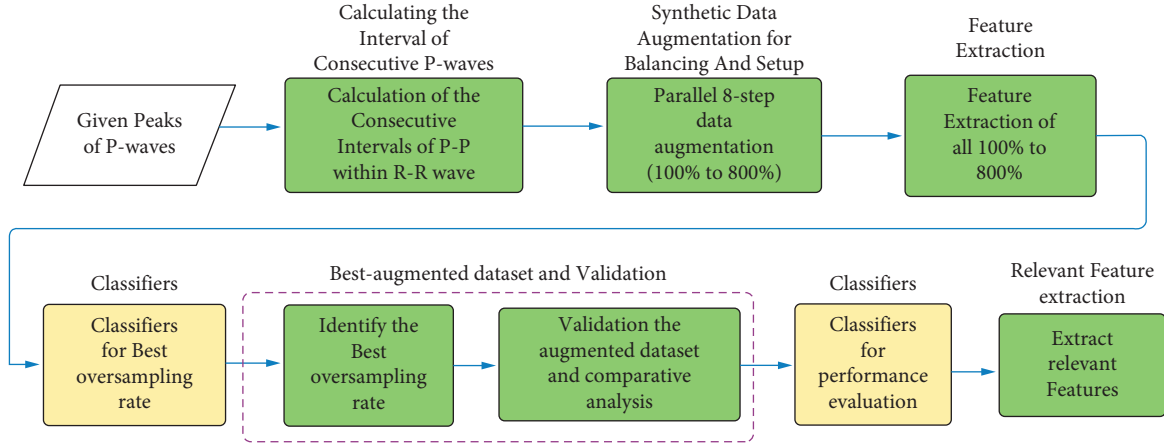


FIGURE 2: Block diagrams of the proposed methodology. In contrast, the novel contribution of this research is highlighted in green shades.

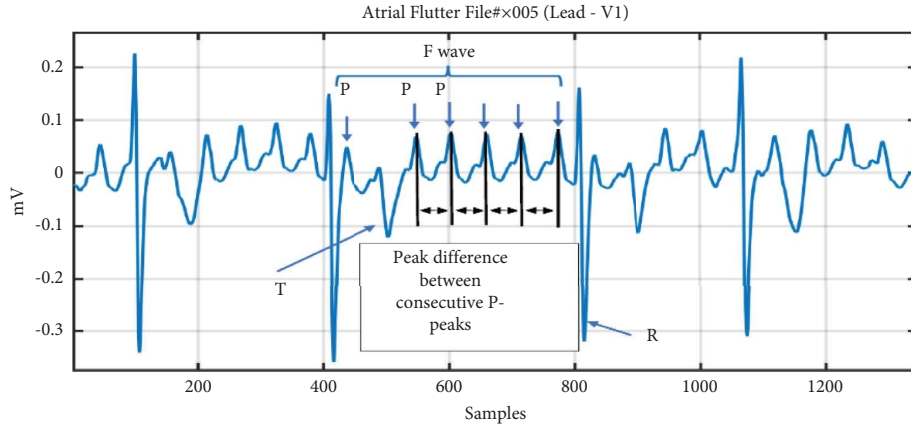


FIGURE 3: Measures of peak-to-peak intervals between consecutive P-waves within R-R waves in AFL with 7:1 ratio.

$$S_{J^{\text{th}}\text{Interval}}^{\text{Modified}} = X_{J^{\text{th}}\text{Interval}} + \frac{3}{2}\alpha(X_{J^{\text{th}}\text{Interval}}^k - X_{J^{\text{th}}\text{Interval}}), \quad (2)$$

where all the terms are as defined previously. The additional coefficient  $3/2$  reexpands the variance of the newly augmented data to match the original dataset. The complete derivation and concept have been derived in Appendix (A.1). The pseudocode for the algorithm of modified-SMOTE is shown in Algorithm 1.

**2.3.3. Smoothed-Bootstrap.** The bootstrap is a conventional method based on resampling with replacement from a conveniently small dataset to construct bootstrap datasets. These derived datasets serve to estimate different functionals of the original distribution. However, the bootstrap distributions are typically discrete. For example, the cumulative distribution function of the classic bootstrap is

$$F_{\text{Bootstrap}}(X) = \sum_{j=1}^{N_B} \frac{\theta(X - X_j)}{N_B}, \quad (3)$$

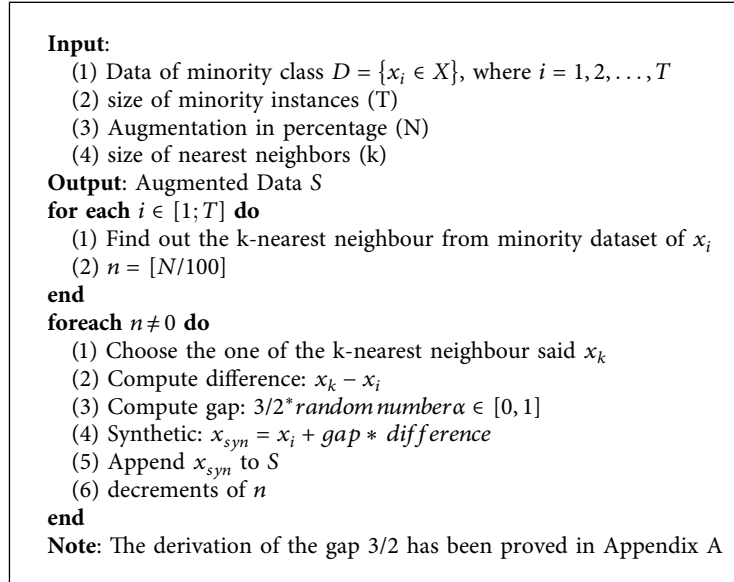
where  $\theta$  is the Heaviside step function or unit step function,  $X$  is the distribution value of the minority class, and  $N_B$  is the size of the bootstrap dataset. Smoothed-bootstrap allows to

smoothen the distribution and renders it continuous. First, the original points are randomly shifted before every resampling [23]. A specially chosen continuous kernel function  $f_j$  is used as follows:

$$F_{\text{Smoothed}}(X) = \sum_{j=1}^{N_B} \frac{f_j(X - X_j)}{N_B}, \quad (4)$$

and thus it renders the distribution continuous. Data points can then be sampled from this distribution. In this research, a nonparametric kernel density function is used.

**2.4. Data Augmentation Setup.** The minority class (FAFL) is used in Figure 4 to generate synthetic data. The interval series of each ECG record were used. For each original interval series, an interval is randomly selected and used to generate a synthetic interval, based on the three algorithms. The process is repeated until the number of synthetic intervals matches that of the current original interval series. This is then repeated for all FAFL ECG records until the desired augmentation rate is achieved (e.g., 200% augmentation rate means each of the five FAFL ECG record generates two synthetic interval series, for a total of 10 synthetic interval series).



ALGORITHM 1: Algorithm of modified-SMOTE.

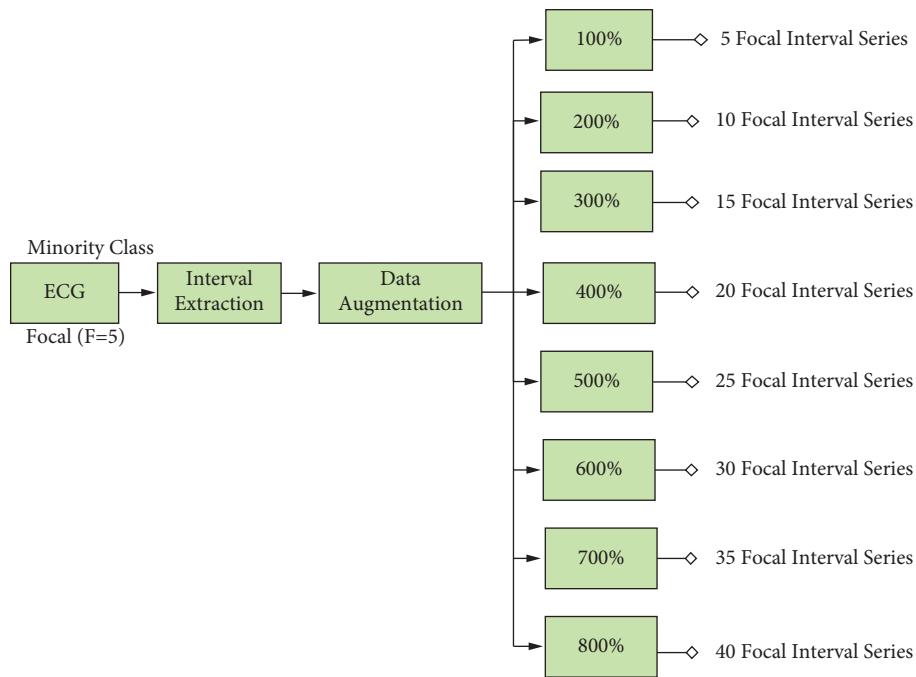


FIGURE 4: Flow of synthetic data generation.

It is known that misuse of data augmentation can lead to biased performance as synthetic samples introduce non-natural qualities (e.g., correlation). It is then a question about how much synthetic data one should introduce. To analyze the effect of augmentation on classifier performance

as well as on the best augmentation rate, this study considers rates from 100% (twice the size of the original dataset) until 800% ( $9 \times$  the size of the original dataset) in steps of 100%. Note that at an 800% oversample rate, the minority class becomes exactly balanced with the majority class.

TABLE 3: Feature list.

		Statistics	Description
F1	Central tendency	Mean	Average value of a set
F2		Median	Middle value of a set
F3		Mode	Most common value of a set
F4	Dispersion	Standard deviation	Spread of the values in a set
F5		Variance	
F6	Shape	Skewness	Distribution asymmetry
F7		Kurtosis	Distribution tailedness
F8	Length (interval)	Maximum	Largest value of a set
F9		Minimum	Smallest value of a set
F10		Sum	Sum of all values of a set

*2.5. Feature Extraction.* The characteristic features describing the interval series have been divided into the following four categories concerning the data: central tendency, dispersion, shape, and length. A total of 10 features were considered and is summarized in Table 3. Due to the difference in mechanism, it is expected that FAFL and MAFL would display different values for these features.

*2.6. Feature Selection Method and Classifiers.* Feature selection aims to avoid issues due to large complexity classifier models (e.g., overfitting and cost ineffectiveness) and improve its performance by selecting only relevant features. We consider in this study two approaches to feature selection, that is, (1) the filter method, using Wilcoxon’s rank-sum test to determine the significant difference in data medians, and (2) the wrapper method, by evaluating all possible feature combinations (1023 combinations) and determining, for each feature, the feature score. The score is determined as follows:

For every combination length (e.g., single feature and pair of features), a score of 1 is assigned to a feature if it was found to participate in the combination with the maximum accuracy, for that particular combination length. The scores for all combination lengths are added and normalized by the number of features available. Relevant features are those whose scores are closest to 1, and vice versa. The algorithm of both the filter method and the wrapper method is shown in the form of pseudocode in Algorithm 2. The wrapper method allows the evaluation of the relevancy on combinations of features and accounts for possible interactions amongst them, in contrast to the filter, which can only evaluate features one by one.

Three linear classification models: Linear Discriminant Analysis (LDA), Logistic Regression (LOG), and Support Vector Machine (SVM) have been used. Note that nonlinear classifiers have not been considered to avoid further issues related to overfitting due to the scarcity of data.

### 3. Results

*3.1. Validation of the Data Augmentation.* The original minority class distribution has been used to generate synthetic and augmented distribution using three different data augmentation techniques. The comparison has been simplified into the following four tests for validation. For each of

these tests, 100 augmented datasets were generated for each augmentation rate (i.e., 100% to 800%), and the test results were averaged. Unless otherwise stated, the augmentation rate is fixed to the best rate, which is 400%. However, the impact of data augmentation from the minority class of five focal ECGs (in terms of intervals) in eight equal steps is demonstrated in Appendix B Figures 5–7.

*3.1.1. Graphical Exploratory Analysis (CDF and Boxplot).* The average empirical cumulative distribution function (CDF) is shown in Figure 8 for all three augmentation techniques. The original empirical CDF is shown for comparison. The bin size was set to 5 ms arbitrarily. All three techniques correctly follow the original pattern. Smoothed-bootstrap follows the original CDF most closely. Modified-SMOTE and SMOTE present some skewness in the range from 180 ms to 205 ms. However, modified-SMOTE presents less skew compared to SMOTE.

Figure 9 shows the average box plot for all three techniques along with its original focal intervals. It has been observed that the central quartile (50%, median) of all three augmentation techniques has the same value as each other and is relatively close to the original dataset. As a quantitative comparison, the median difference in quartile values between the original and each of the three augmented datasets is calculated and summarized in Table 4. The difference in all augmented dataset medians with the original is very small. However, modified-SMOTE has been found to match the original regarding upper and lower quartile ranges (75% and 25%). This shows that modified-SMOTE is a better technique.

*3.1.2. Nongraphical Exploratory Analysis (Descriptive Statistics and the Goodness-of-Fit Test).* The Kolmogorov F02D Smirnov goodness-of-fit test has been used to measure the degree of disagreement between the empirical CDFs of the original and augmented datasets. The p value of this test was taken as a measure of similarity (higher values theoretically mean higher similarity). Figure 10 summarizes the statistics of the p values. Modified-SMOTE had the largest p value ( $0.64 \pm 0.12$ ), suggesting a very high distribution similarity to the original. SMOTE and smoothed-bootstrap have significantly smaller values than this, with smoothed-bootstrap being the smallest (average  $0.26 \pm 0.07$  vs.  $0.16 \pm 0.09$ ).

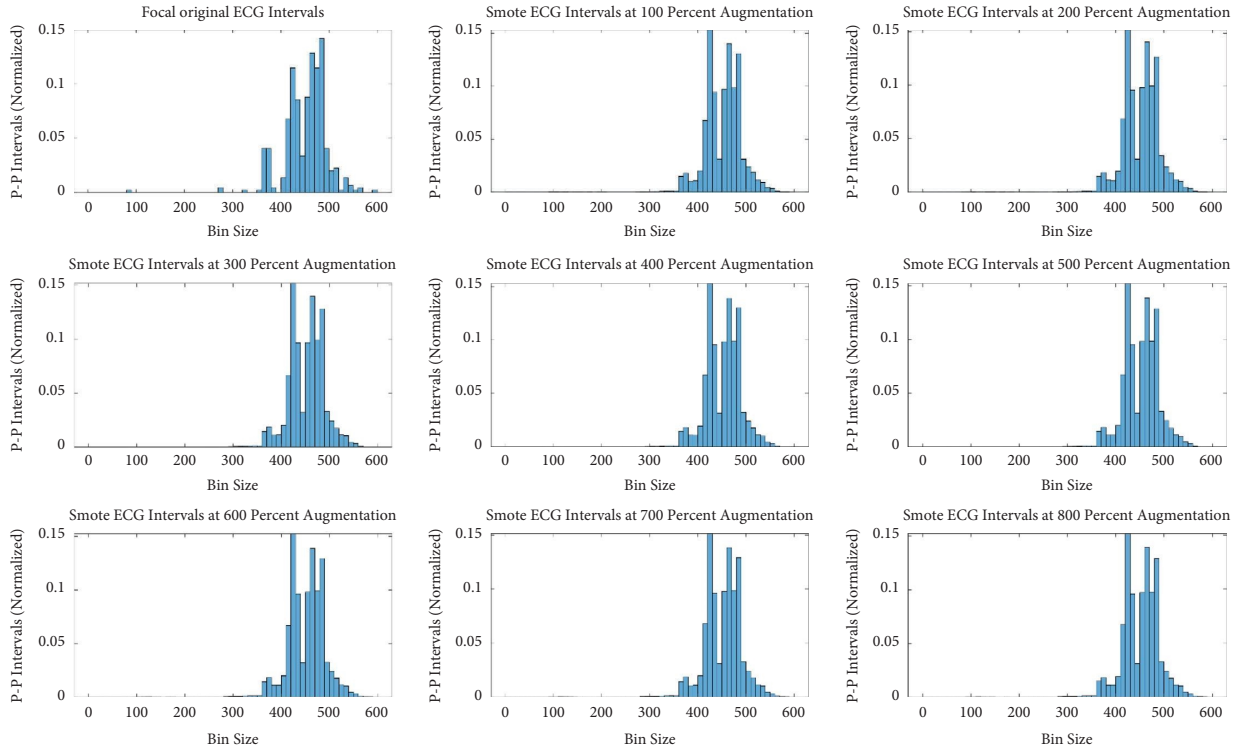


FIGURE 5: Augmented minority (focal) ECG by SMOTE.

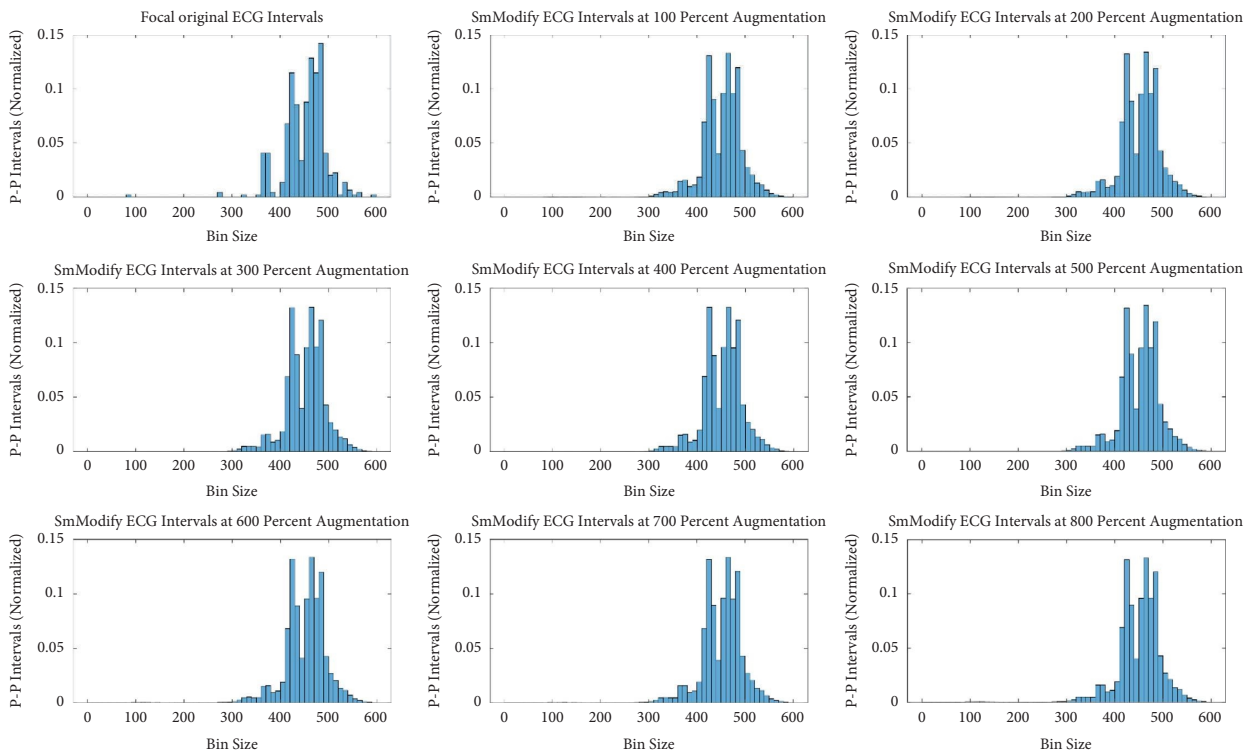


FIGURE 6: Augmented minority (focal) ECG by modified-SMOTE synthetic technique.

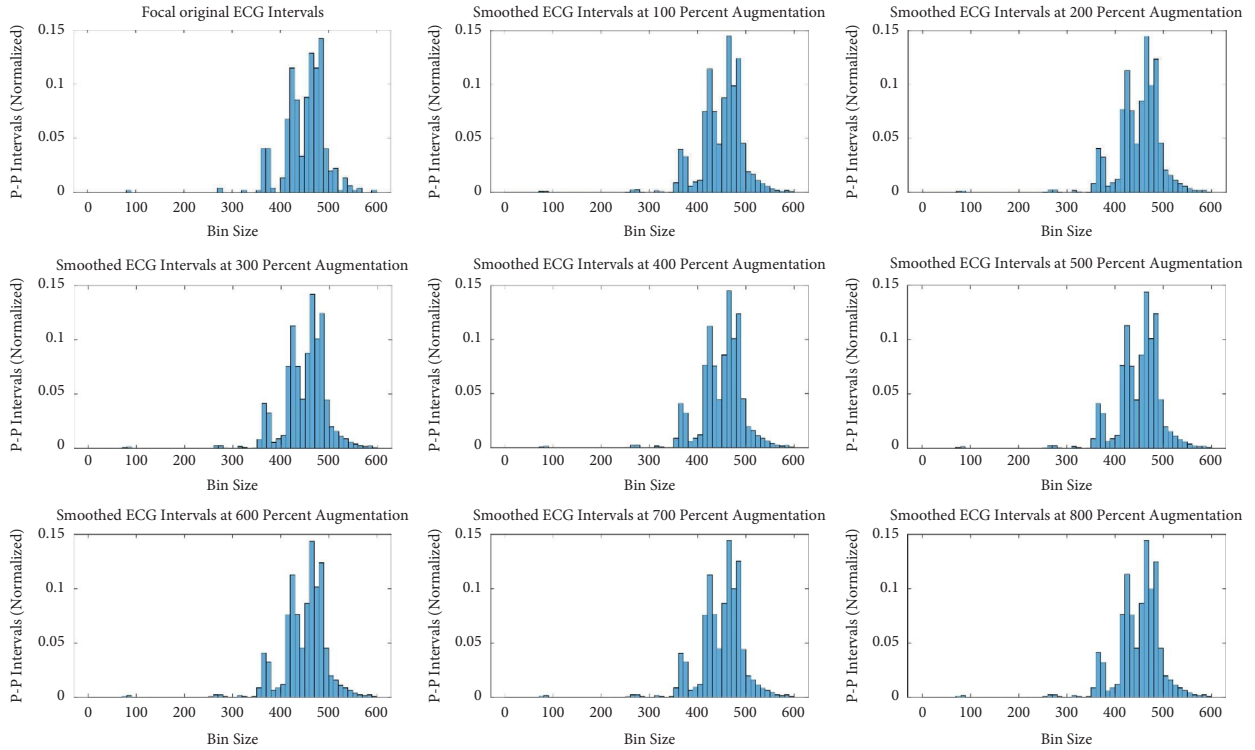


FIGURE 7: Augmented minority (focal) ECG by smoothed-bootstrap synthetic technique.

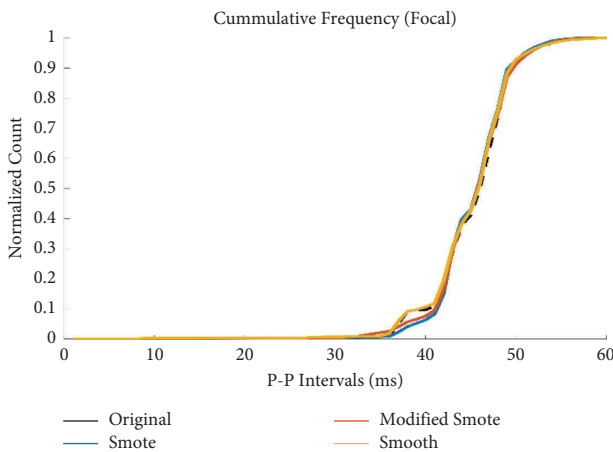


FIGURE 8: Empirical CDF of the original and augmented P-P intervals of all FAFL records at best augmentation rate (400%).

Finally, three descriptive statistics, namely, (1) mean, (2) variance, and (3) skewness of the augmented dataset were compared to the original. The difference between original and augmented dataset statistics was calculated and shown in Table 5 as percentages of the original value. The minimum differences are marked in bold font. It can be observed that smoothed-bootstrap has replicated the closest variance and skewness to the original dataset, whereas modified-SMOTE has the minimum difference in the mean only.

**3.2. Selection of the Best Augmentation Rate.** Figure 11 shows the maximum performance in terms of specificity of the

three classifiers, for each feature combination length. The ideal rate would be the one that maximizes accuracy, sensitivity, and specificity (considering MAFL as the target class). However, under difficult conditions such as heavy class imbalance, these measures have to be considered carefully. Since the imbalance here affects the negative class (i.e., FAFL), we propose to trade off better specificity against lower sensitivity.

To identify the best augmentation rate, the average curve of all augmentation rates was calculated for each feature combination length. The augmentation rate, whose curve has the minimum overall distance from the average curve was chosen as the best augmentation rate. The best rate is shown in red dashed lines in Figure 11 (referring to the right axis). It is observed that the best rate is not uniform across the different combination lengths. It is primarily stable between 400% and 600%. We have selected a lower range of 400% as the best augmentation rate to prefer the minimum synthetic ratio compared to 500% and 600%.

**3.3. Performance Evaluation by Linear Classifiers for Best Oversampling Rate.** Three linear classifiers LDA, LOG, and SVM, and their performances are shown in Figures 11 and 12. One common behavior is that the specificity has increased with an imbalance reduction between FAFL and MAFL. Contrarily, sensitivity, and accuracy decreased with increasing augmentation rate.

The maximum performance at 400% augmentation is summarized in Table 6 and shown in Figure 12. The LOG classifier has the highest performance among the linear classifiers, with mean values of 76.13%, 41.42%, and 93.76% for

```

Data: Set of features  $\mathcal{F} = \{f_1, f_2, \dots, f_l\}$ 
Set all score counters  $s_{ij}, i \in [1; I]$  to 0 and  $j \in \text{NO}$  of augmentation techniques ( $N$ )
foreach  $j \in [1; N]$  do
  **feature scoring wrapper approach**
  foreach  $l \in [1; L]$  do
    Set  $\epsilon \neq \emptyset$ 
    Perform a wrapper evaluation of all possible
      feature combination of length  $l$  and
      calculate the performance metric  $S$ 
    Determine the features participating in the best
      combinations of length  $l$  and store them  $\epsilon$ 
    foreach  $f_i \in \epsilon$  do
      Increment  $s_{ij}$ 
    end
  end
foreach  $i \in [1; I]$  do
   $s_{ij} = s_{ij}/I$ 
end
****feature scoring for filter method****
Set  $\epsilon_a = \alpha$ 
foreach  $\mathcal{F} \in [1; M]$  do
  Determine the feature which lies within the  $\alpha$  and store them in  $\epsilon_a$ 
end
end

```

ALGORITHM 2: Feature scoring algorithm.

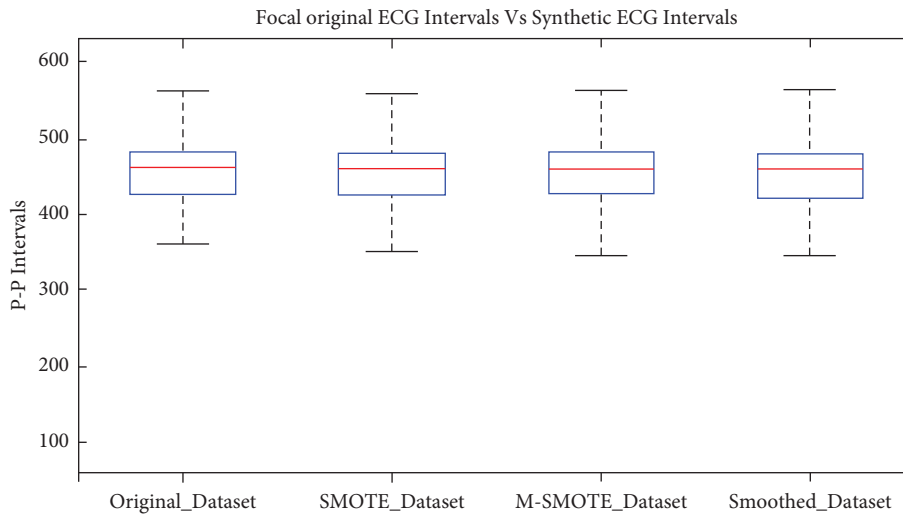


FIGURE 9: Comparison of 400 percent augmented focal.

TABLE 4: Difference of original dataset from augmented dataset at 400% augmentation rate.

Quartile (%)	SMOTE	Modified-SMOTE	Smoothed-bootstrap
25	0	0	6
50	1.5	1.5	1.5
75	2	0	3

accuracy, specificity, and sensitivity, respectively, and for all three augmentation techniques. The LDA classifier has the maximum performance under modified-SMOTE, with 73.86%, 29.51%, and 95.20% accuracy, specificity, and sensitivity,

respectively. Classifier sensitivity has improved after augmentation. The improvement in performance shows that minimizing the imbalance has a positive impact on performance.

**3.4. Relevant Feature Selection Methods.** Two kinds of supervised feature selection methods have been used for identifying the relevant features that can differentiate MAFL from FAFL. The Wilcoxon rank-sum  $p$  values (filter method) and the feature scores (wrapper approach) are shown in Tables 7 and 8 for the original and best-augmentation rate, respectively. Relevant (grey

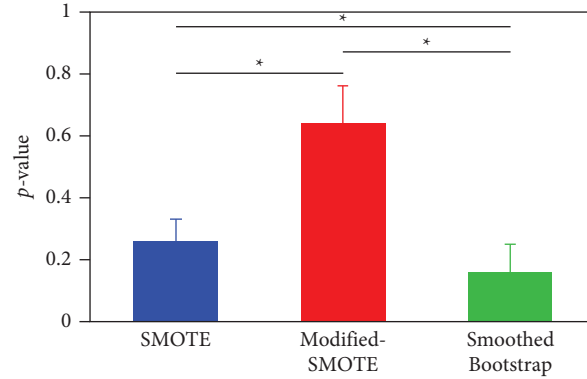


FIGURE 10: Comparison of Kolmogorov F02D Smirnov test p values at the best augmentation rate. Asterisks indicate significant differences between augmented datasets ( $t$ -test,  $pp < 0.05$ ).

TABLE 5: Comparison in the percentage of the augmented dataset by descriptive statistics on the minority data (focal AFL).

	Mean			Variance			Skewness		
	SMOTE	Modified	Smoothed-bootstrap	SMOTE	Modified	Smoothed-bootstrap	SMOTE	Modified	Smoothed-bootstrap
	Difference between statistics of the original dataset and the augmented dataset (in percentage)								
100	-0.19	<b>-0.17</b>	0.43	25.52	10.65	<b>-4.36</b>	18.34	19.98	<b>3.45</b>
200	-0.19	<b>-0.14</b>	0.43	25.74	9.69	<b>-4.27</b>	23.87	19.74	<b>4.69</b>
300	-0.18	<b>-0.17</b>	0.42	25.61	10.59	<b>-4.01</b>	20.70	18.82	<b>4.43</b>
400	-0.20	<b>-0.16</b>	0.41	25.93	10.59	<b>-3.83</b>	20.35	18.92	<b>4.85</b>
500	-0.19	<b>-0.17</b>	0.44	25.95	9.95	<b>-4.26</b>	21.51	20.11	<b>4.51</b>
600	-0.20	<b>-0.15</b>	0.41	25.93	10.26	<b>-3.78</b>	21.53	20.01	<b>4.81</b>
700	-0.19	<b>-0.17</b>	0.42	25.45	10.77	<b>-3.97</b>	21.00	20.14	<b>4.42</b>
800	-0.19	<b>-0.17</b>	0.41	25.85	10.55	<b>-4.04</b>	22.86	20.80	<b>5.15</b>

There are three augmentation techniques (Smote, Modified-Smote, and Smoothed-Bootstrap) for each parameter (Mean, Variance, and Skewness). The technique with the minimum difference is indicated in bold font.

highlighted) features can be seen to have high feature scores. Despite the  $p$  values being nonsignificant for most features in general, it can be seen that the wrapper can highlight features that most probably perform better when combined.

#### 4. Discussion

The 12-lead surface ECG remains a staple tool for heart disease diagnosis. However, it is rarely used for AFL mechanism diagnosis. On the other hand, it is widely used to distinguish AFL from atrial fibrillation [6]. The proposed methodology thus represents a contribution in the use of 12-lead ECG as a tool for AFL mechanism discrimination. This in turn allows clinicians to have an early insight into the ablation strategy, thus economizing time and resources.

Three classifiers LDA, LOG, and SVM, have been used to evaluate the performance of the original dataset. The accuracy, specificity, and sensitivity obtained are 91%, 17%, and 100%, respectively, for the logistic regression classifier, with similar results for two other linear classifiers. It has been observed that the specificity in all three classifiers exhibited abysmal performance ( $< 20\%$ ). This was caused by the heavy class imbalance where there was a 1:8 ratio of focal AFL to macroreentry AFL data record. Classifier bias on the majority class cannot be avoided. Therefore, data augmentation techniques were used to minimize the imbalance. At 400%

augmentation, the LOG classifier achieved an accuracy, specificity, and sensitivity of 76%, 40%, and 94%, respectively.

The augmentation here does not serve to “improve” the classifier performance: rather, the focus was to “regularize” the obvious bias due to imbalance. It can be seen that despite a drop in overall accuracy, sensitivity did not drop significantly and yet specificity increased more than twice. This suggests that the use of augmentation helps to estimate the correct performance in regard to classifying focal AFL.

*4.1. Validation of the Data Augmentation.* The results of all four tests of data validation are summarized in Table 9. It is initially difficult to identify the uniformly best technique since all three techniques have competing ranks. However, modified-SMOTE has never been listed as rank 3. Therefore, it can be suggested that modified-SMOTE is overall a better technique for data augmentation among the proposed three techniques.

It can be seen in Table 5 that modified-SMOTE produces datasets with generally less variance compared to SMOTE, as suggested in Appendix A. The issue of variance shrinkage was highlighted by Blagus and Lusa [17], but no solution was proposed. Our original contribution here produces a general tool for generating a new dataset with similar first-order and



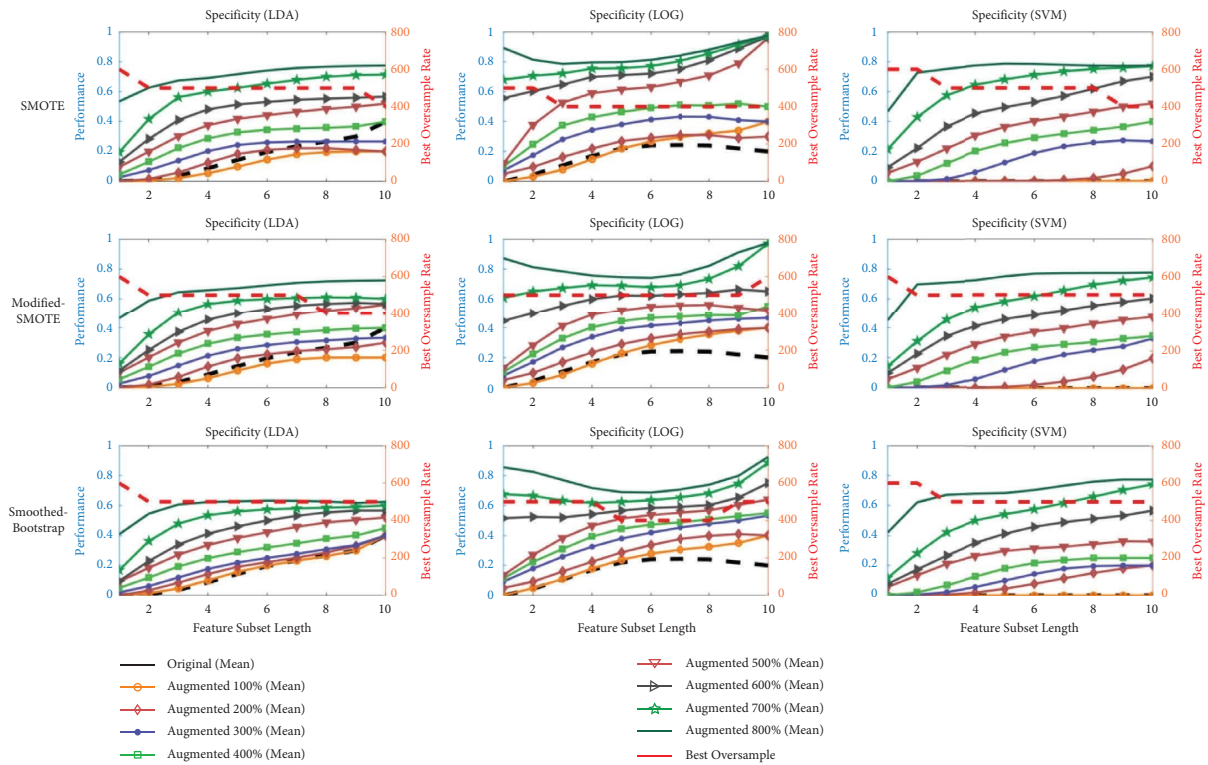


FIGURE 11: Specificity of proposed augmentation techniques by three classifiers (LDA, LOG, and SVM) for all feature subset lengths. Right y-axis indicates performance of the classifier using synthetic samples, and left y-axis shows the best oversample rate for a particular feature-length (red line).

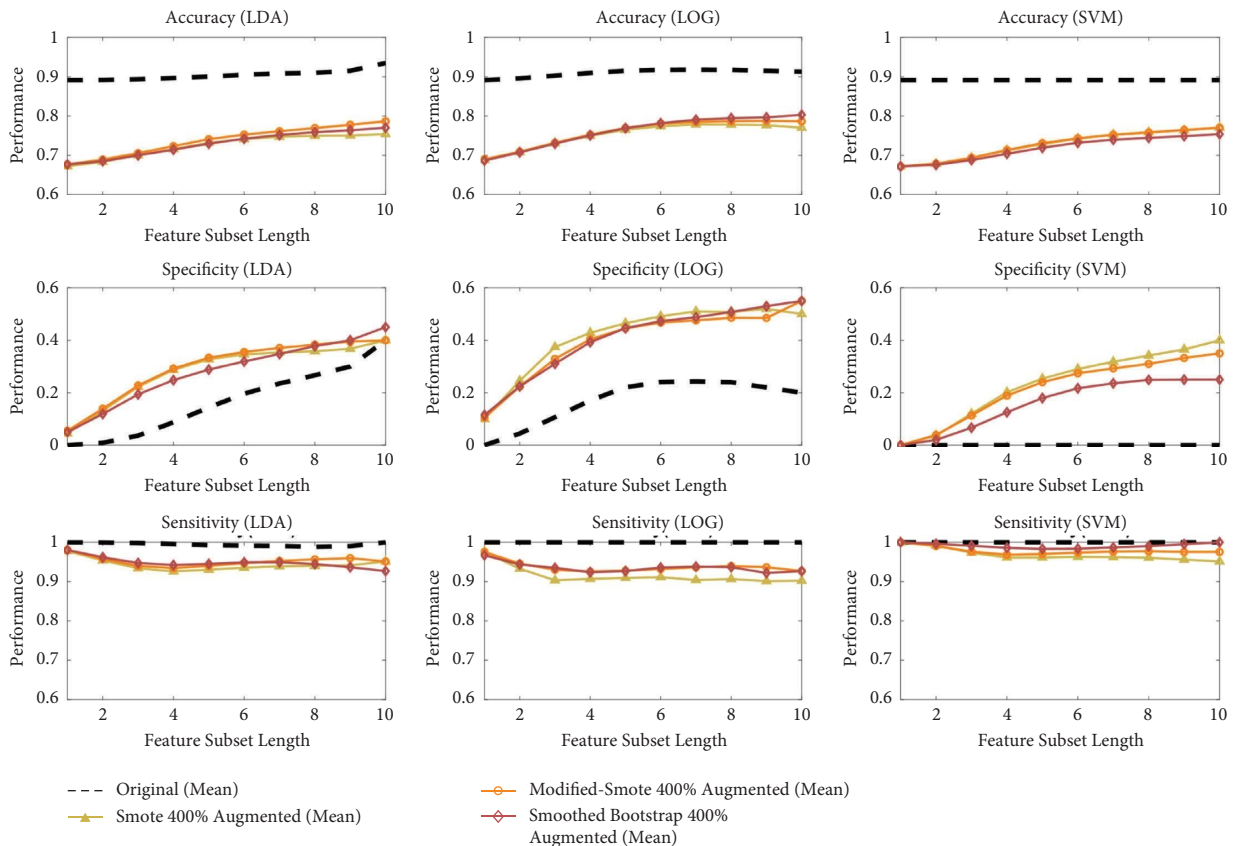


FIGURE 12: Performance evaluation of the proposed augmentation techniques by three classifiers (LDA, LOG, and SVM) for all feature subset lengths.

TABLE 6: Mean value of classifier performance at 400% data augmentation.

Technique	Performance	LDA	LOG	SVM
SMOTE	Accuracy	72.48	75.22	72.76
	Specificity	28.38	41.42	23.30
	Sensitivity	94.29	91.54	96.81
Modified-SMOTE	Accuracy	73.86	75.82	72.89
	Specificity	29.51	39.70	21.41
	Sensitivity	95.20	93.76	97.85
Smoothed-bootstrap	Accuracy	72.96	76.13	71.80
	Specificity	27.94	40.36	15.93
	Sensitivity	94.83	93.56	99.13

TABLE 7: Feature selection methods (both filter and wrapper) at the original dataset.

Features	Original dataset			
	Filter method	Wrapper method		
	p value	LDA	LOG	SVM
Mean	0.80	0.9	1	1
Median	1	1	1	1
Mode	0.97	0.9	1	1
Std	0.57	0.9	1	1
Var	0.57	0.9	1	1
Skew	0.34	0.9	0.9	1
Kurt	0.62	0.8	0.7	1
PDmin	0.39	1	1	1
PDmax	0.69	0.7	0.8	1
Interval	0.28	0.8	0.9	1

Note. In the filter method, we used the Wilcoxon rank-sum test.

second-order moments to the original dataset. It can be helpful for other researchers in handling imbalanced datasets, which is a real problem in the biomedical field.

**4.2. Selection of Relevant Features.** The relevant features that differentiate the mechanism of AFL with the highest performance at 400% augmentation have been extracted from two different feature selection methods. The results of the filter method and the wrapper approach are already shown in Table 8. It is challenging to decide the relevant feature with a single method since many feature subsets have scored more significantly than the arbitrary threshold of 0.8. Contrarily, it is simple to decide the relevant features in the filter method as only one is significant. However, this method compares single features and ignores dependencies.

One solution is to compare the two methods to conclude on feature relevance. According to this, three relevant features are highlighted (in order of decreasing relevance): F10: sum of all consecutive intervals, F8: the minimum difference between consecutive P-wave intervals, and F5: variance.

A peak-to-peak interval of two consecutive P-waves contains two temporal information, that is, the P-wave duration, defined as the time length from its onset until its end, and the isoelectric line duration. Hence, there is an influence of P-wave morphology in our measured P-P interval. Therefore, the variation in both P-wave and isoelectric line duration, due to conduction path variability, contributes both to the differentiation of focal and

macroreentrant AFL. The sum of all consecutive intervals (F10) has been selected as the relevant feature based on this hypothetical phenomenon, and it has been found to be different for focal AFL from macroreentrant AFL ( $22.77 \pm 12.13$  vs.  $14.15 \pm 10.57$ , respectively,  $p < 0.05$ ). Furthermore, the minimum peak-to-peak interval length (F8) discriminated the AFL mechanism based on fast and slow conduction velocity ( $393.68 \pm 54.51$  vs.  $424.63 \pm 74.48$ , respectively).

In summary, the sum of all consecutive intervals is the relevant feature to discriminate the mechanism. Finally, the acceptance of our study's hypothesis about the flutter mechanism has also led to the conclusion that the variable feature is a more relevant subset for distinguishing the atrial flutter mechanism.

**4.3. Performance Evaluation by Linear Classifiers.** It has been concluded from the previous section that the best-augmented ratio is 400%, and the modified-SMOTE is the appropriate technique for augmentation. Therefore, the performance of the proposed method has been conducted at 400% of the modified-SMOTE by using five-fold cross-validation. Its results are shown in Table 10 concerning accuracy, specificity, and sensitivity. These results validate that the consecutive intervals of P-P peaks are the significant factors for the discrimination of the AFL mechanism from 12-lead surface ECG.

TABLE 8: Median and standard deviation values of both filter and feature selection methods on all synthetic techniques at 400% data augmentation.

Features	Median of p value by filter method				Median of feature scores by wrapper method					
	SMOTE	Modified-SMOTE	Smoothed-bootstrap	Smoothed-bootstrap	LDA	LOG	SVM	LDA	LOG	SVM
Mean	0.54	0.55	0.55	0.55	0.7	0.7	0.8	0.6	0.7	0.7
Median	0.88	0.91	0.84	0.84	0.7	0.7	0.8	0.5	0.6	0.7
Mode	0.95	0.94	0.87	0.87	0.5	0.7	0.7	0.4	0.7	0.7
Std	0.86	0.40	0.34	0.34	0.7	0.6	0.8	0.7	0.7	0.8
Var	0.86	0.40	0.34	0.34	0.7	0.6	<b>0.8</b>	<b>0.8</b>	0.7	<b>0.8</b>
Skew	0.11	0.11	0.28	0.28	0.6	0.7	0.8	0.6	0.7	0.7
Kurt	0.51	0.74	0.76	0.76	0.7	0.8	0.9	0.7	0.7	0.7
PDmin	<b>0.22</b>	<b>0.15</b>	<b>0.13</b>	<b>0.13</b>	0.7	0.7	<b>0.9</b>	<b>0.8</b>	<b>0.9</b>	<b>0.8</b>
PDmax	0.27	0.58	0.46	0.46	0.5	0.6	0.7	0.6	0.6	0.7
Interval	<b>0.01</b>	<b>0.01</b>	<b>0.01</b>	<b>0.01</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>
Standard deviation of p value by filter method										
Mean	0.055	0.068	0.109	0.109	0.114	0.132	0.082	0.113	0.095	0.116
Median	0.051	0.071	0.095	0.095	0.110	0.088	0.089	0.107	0.096	0.090
Mode	0.049	0.068	0.091	0.091	0.137	0.090	0.110	0.117	0.123	0.138
Std	0.078	0.085	0.146	0.146	0.089	0.122	0.116	0.122	0.130	0.116
Var	0.078	0.085	0.146	0.146	0.077	0.149	<b>0.084</b>	<b>0.064</b>	0.133	<b>0.085</b>
Skew	0.071	0.094	0.148	0.148	0.125	0.124	0.177	0.100	0.082	0.160
Kurt	0.188	0.180	0.188	0.188	0.123	0.176	0.102	0.122	0.141	0.089
PDmin	<b>0.061</b>	<b>0.038</b>	<b>0.056</b>	<b>0.056</b>	0.112	0.189	<b>0.169</b>	<b>0.115</b>	<b>0.034</b>	<b>0.145</b>
PDmax	0.053	0.089	0.074	0.074	0.167	0.098	0.095	0.105	0.115	0.093
Interval	<b>0.001</b>	<b>0.00</b>	<b>0.002</b>	<b>0.002</b>	<b>0.010</b>	<b>0.031</b>	<b>0.010</b>	<b>0.026</b>	<b>0.038</b>	<b>0.024</b>
Standard deviation of feature scores by wrapper method										
Mean	0.055	0.068	0.109	0.109	0.114	0.132	0.082	0.113	0.095	0.116
Median	0.051	0.071	0.095	0.095	0.110	0.088	0.089	0.107	0.096	0.090
Mode	0.049	0.068	0.091	0.091	0.137	0.090	0.110	0.117	0.123	0.138
Std	0.078	0.085	0.146	0.146	0.089	0.122	0.116	0.122	0.130	0.116
Var	0.078	0.085	0.146	0.146	0.077	0.149	<b>0.084</b>	<b>0.064</b>	0.133	<b>0.085</b>
Skew	0.071	0.094	0.148	0.148	0.125	0.124	0.177	0.100	0.082	0.160
Kurt	0.188	0.180	0.188	0.188	0.123	0.176	0.102	0.122	0.141	0.089
PDmin	<b>0.061</b>	<b>0.038</b>	<b>0.056</b>	<b>0.056</b>	0.112	0.189	<b>0.169</b>	<b>0.115</b>	<b>0.034</b>	<b>0.145</b>
PDmax	0.053	0.089	0.074	0.074	0.167	0.098	0.095	0.105	0.115	0.093
Interval	<b>0.001</b>	<b>0.00</b>	<b>0.002</b>	<b>0.002</b>	<b>0.010</b>	<b>0.031</b>	<b>0.010</b>	<b>0.026</b>	<b>0.038</b>	<b>0.024</b>

The relevant features have been emphasized, and the use of bold font indicates that certain sets of features have achieved greater significance than the arbitrary threshold of 0.8.

TABLE 9: Summary of data validation results at best augmentation rate.

Parameter	Subparameter	Rank-1	Rank-2	Rank-3
CDF		Smoothed-bootstrap	Modified-SMOTE	SMOTE
Boxplot		Modified-SMOTE	SMOTE	Smoothed-bootstrap
KS test		Modified-SMOTE	SMOTE	Smoothed-bootstrap
Statistical	Mean	Modified-SMOTE	SMOTE	Smoothed-bootstrap
	Var	Smoothed-bootstrap	Modified-SMOTE	SMOTE
	Skew	Smoothed-bootstrap	Modified-SMOTE	SMOTE

TABLE 10: Performance evaluation by linear classifiers at 400% data augmentation with five-fold cross-validation.

Technique	Performance	LDA	LOG	SVM
Modified-SMOTE	Accuracy	77.81	76.88	77.45
	Specificity	41.35	49.50	36.25
	Sensitivity	95.60	90.24	97.56

**4.4. Comparative Analysis.** The definitions of AFLs and a new classification correlated with mechanisms were proposed in 2001 by an international panel of specialists [3]. They have briefly explained the tachycardia mechanism concerning mapping, transient entrainment, and ECG pattern. According to them, during AT, the presence of isoelectric lines indicates the presence of underlying focal mechanisms in a vast majority of patients. In contrast, the lack of isoelectric lines indicates the presence of macroreentry mechanisms in a vast majority of patients during AT. Importantly, it is also possible to observe isoelectric lines in macroreentry, however, only if significant atrial scarring is present. This statement paves the way for researchers in focal and macroreentrant atrial flutter cases. An extensively wide study is focused on isoelectric intervals in discriminating the focal from macroreentrant by invasive and noninvasive procedures. However, limited research was found in non-invasive mechanisms for discrimination of the AFL mechanism, especially the 12-lead surface ECG, discussed here and it is summarized in Table 11.

Two methods for discriminating focal from macroreentrant atrial flutter were proposed by Brown et al. [8]. First, the P-wave duration of the focal should be less than 160 ms (accuracy, 80%), and second, the ratio of P-wave duration to tachycardia cycle length should be less than 45% (accuracy, 95%). This model is highly dependent on the delineation of atrial activities to calculate the cycle length of each P-wave, which requires a high signal processing technique. In contrast, our results were measured from the proposed consecutive P-waves, which were measured from the atrial activity peaks within the R-R intervals. Therefore, the proposed model keeps safe from the advanced signal processing and discriminates the atrial flutter mechanism without requiring the delineation of atrial activities (specificity, 76.88%). Moreover, the original dataset percentage ratio between macroreentrant and focal was 65:35, whereas in our study case, it was 89:11. After augmentation it became 67:33.

Three relevant features were extracted by Chang et al. [10] based on PWM (P-wave morphology): lower voltage in macroreentrant as compared with focal ( $1.3 \pm 0.3$  vs.  $1.5 \pm 0.2$  mV,  $pp = 0.02$ ); high incidence of the positive

polarity of lead V6 in focal (88% vs. 55%,  $p = 0.03$ ); and longer cycle length in focal ( $296 \pm 107$  vs.  $224 \pm 25$  ms,  $p = 0.01$ ). This case was performed experimentally based on a retrospective analysis and, similarly, required advanced signal processing for morphological analysis of atrial activities. In contrast, the proposed model is directly independent of the delineation and morphology of atrial activities. However, our proposed model includes the cycle length of atrial activity and isoelectric interval without advanced signal processing into the measured consecutive P-P interval within the R-R interval. In detail, the consecutive P-P interval has three pieces of information, such as (i) the approximate second half cycle of the first P-wave, (ii) is the isoelectric interval between consecutive p-waves, and (iii) approximately the first half cycle of the second P-wave. We identified the sum of all consecutive intervals (F10) as a relevant feature extracted from our proposed model, which discriminates between focal and macroreentrant AFL ( $22.77 \pm 12.13$  vs.  $14.15 \pm 10.57$ ), respectively.

Recently, a study generated synthetic datasets through eight torso models using twenty different original AFL mechanisms [11, 12]. They produced 1,256 sets of 12-lead ECG records through a forwarding solution. Furthermore, six RQA-based characteristics were retrieved using two approaches, revealing that a 12-lead surface ECG can characterize the differentiation between FAFL and MAFL. With this in mind, we have generated synthetic ECG intervals from the available minority dataset, which contains non-CTI-based left and right circuits. We generated ECGs in the feature space (consecutive P-P interval) instead of the standard time domain because oversampling techniques used in this model worked in feature space. Our results show that at the best-oversampling rate, the minimum P-P interval length (F8) discriminates the AFL mechanism based on fast and slow conduction velocity ( $393.68 \pm 54.54$  vs.  $424.63 \pm 74.48$ , respectively).

**4.5. Limitation and Future Works.** This study is based on the variation of intervals between two consecutive atrial activities. At least two atrial activities must be visibly present between the ventricle activity in the ECG. In terms of ratio, it can also be said that the ratio of atrial and ventricle activity must be greater than 2:1. This criterion is oftentimes strict and renders the data collection a burdensome task. The selection of a maximum delay between the two P-waves is based on assumptions. It should ideally be set after consultation with several electrophysiologists since slower rates can be observed. Further clinical data should be added to handle the imbalance issue in the study dataset.

TABLE 11: Related works on discrimination of AFL mechanism (PWM: P-wave morphology, CL: cycle length, and NA: not applicable).

Author	Year	Ratio macro/focal (ECGs)	Technique	Classifier	Parameter (significant)	Performance
Brown et al.	2007	27/14	PWM, CL	NA (autocorrelation performed)	PP < 160 ms P/CL < 45%	P: Sensitivity: 90% Specificity: 90% P/CL: Sensitivity: 86% Specificity: 98% Accuracy focal: 93%
Chang et al.	2011	51/17	PWM CL	NA (empirical-based study)	V6 > 0.9 mV, CL > 265 ms V6 < 0.9 mV, CL > 290 ms	Accuracy macro: 88% Accuracy focal: 100% Accuracy macro: 100%
Luongo et al.	2020	Original: 11/9 Augmented: 1256	Recurrence quantification	Decision tree (DT), KNN, and radial basis neural network (rbNN)	RQA-based features	Hit rate: 67.7%
Proposed model		Original: 41/5 Augmented: 15 (focal)	Consecutive P-P intervals (within R-R interval)	LDA, LOG, and SVM	Sum of all consecutive intervals ( $p < 0.05$ )	Accuracy: 76.88% Specificity: 49.50 Sensitivity: 90.24%

In this study, the modification of the SMOTE algorithm for correcting variance shrinkage was performed, assuming that the random multiplier  $\alpha$  was drawn from a uniform distribution of the modified range. This was proven to theoretically preserve the original moments of the data distribution up to the second moment. It is an open question about what other distributions may be considered to preserve other data properties. This research can also be extended by exploring more valuable classifiers after balancing the dataset with new samples as future work.

The proposed modified-SMOTE was evaluated on two different mechanisms to validate the modification. First, we have theoretically proven the concept of the modification of the classical SMOTE. Then, we performed a comparative analysis of the modified algorithm, classical algorithm, and other oversampling techniques on the real dataset. However, the performance of modified-SMOTE should be analyzed on public datasets to compare and validate its results with other kinds of SMOTE modification.

## 5. Conclusion

This noninvasive study helps identify the AFL mechanism using 12-lead surface ECG, which allows insight into the disease before the catheter ablation procedure. Consecutive intervals of P-waves are hypothesized to contain crucial information regarding the AFL mechanism. Our findings indicate that they are helpful in the discrimination of focal AFL and macroreentry AFL, which does not rely on advanced signal processing such as the measure of the delineation, onset, and offset of the atrial waves. This study has also applied several data augmentation strategies to cure class imbalance in the original dataset. Based on a classical algorithm, a novel augmentation method (modified-SMOTE) was modified to correct a theoretical issue present in the original algorithm.

Three linear classifiers have been used to discriminate against the AFL mechanism. At the best augmentation rate of 400%, the logistic regression classifier achieved an average sensitivity, specificity, and accuracy of 90.24%, 49.5%, and 76.88%, respectively. It was concluded that the sum of all consecutive atrial activities is a relevant feature to differentiate the AFL mechanism.

## Appendix

### A. Proof of Synthetic Sample Variance Shrinkage in SMOTE and a Countermeasure

Let  $X_j, X_j^k \in R^M$  be two  $k$ -near samples from a dataset  $\chi$  of  $N$  samples. According to SMOTE [15], a new synthetic sample  $S_j$  is generated according to the following rule:

$$S_j = X_j + \alpha(X_j^k - X_j), \quad (\text{A.1})$$

with  $\alpha$  as a random number from the standard uniform distribution.

The expression in (A.1) can be rearranged as follows:

$$\begin{aligned} S_j &= X_j + \alpha(X_j^k - X_j) \\ &= X_j + \alpha X_j - \alpha X_j^k \\ &= (1 - \alpha)X_j + \alpha X_j^k, \end{aligned} \quad (\text{A.2})$$

which makes it clear that the synthetic sample is an interpolation between  $X_j$  and  $X_j^k$ . In light of the bootstrap sampling theory [24],  $\alpha$  can be understood as a ‘‘smoothing’’ factor, allowing the selection of new samples that are not directly within  $\chi$ . A crucial part of this procedure is to preserve the properties of the probability distribution, such that new samples  $S$  have the same first-order (mean) and second-order central (variance) moments.

In what follows, we consider each  $X$  as a realization of a random variable  $\xi$  distributed according to some probability distribution. We then consider the following equation:

$$\sigma_j = (1 - A)\xi_j + \xi_j^k, \quad (\text{A.3})$$

with  $\xi_j, \xi_j^k$  i.i.d., and  $A \sim U(0, 1)$ . Note that this is related to (A.1) when we assume  $X_j, X_j^k$ , and  $\alpha$  are the realization of the respective random numbers, then  $S_j$  can thus be considered an application of the values of the previous. The implication from this consideration is that all  $x \in \chi$  are distributed according to a common probability distribution. In addition,  $\xi_j^k$  could be any one neighbor of  $\xi_j$ . Hence, in such a setup, we consider all samples to be nearest neighbors (i.e.,  $k = N - 1$ ).

The form in (A.3) resembles a mixture distribution with uniformly distributed mixing weights. Assuming that  $\xi_j, \xi_j^k$ , and  $A$  are all independent, we have  $E[\sigma_j] = E[\xi_j]$  and  $\text{Var}[\sigma_j] = (2/3)\text{Var}[\xi_j]$

*Proof.* We resort to the law of total expectation which is

$$\begin{aligned} E[\sigma_j] &= E_A[E_\xi[(1 - A)\xi_j + A\xi_j^k | \alpha]] \\ &= E_A[E_\xi[(1 - A)\xi_j + A\xi_j^k]] \\ &= E_A[(1 - A)E_\xi[\xi_j] + AE_\xi[\xi_j^k]] \\ &= E_A[(1 - A)E_\xi[\xi_j] + E_A[A]E_\xi[\xi_j^k]] \\ &= E_A[1] - E_A[A]E_\xi[\xi] + E_A[A]E_\xi[\xi] = E[\xi]. \end{aligned} \quad (\text{A.4})$$

Then, we resort to the law of total variance, that is,

$$\begin{aligned}
\text{Var}[\sigma_j] &= \text{Var}[E_k[\sigma_j | \alpha]] \\
&= \text{Var}[E[\xi]] + E_A[\text{Var}[(1-A)\xi_j + A\xi_j^k]] \\
&= 0 + E_A[(1-A)^2\text{Var}[\xi_j] + A^2\text{Var}[\xi_j^k]] \\
&= \text{Var}[\xi]E_A[(1-A)^2 + A^2] \\
&= \text{Var}[\xi]E_A[1 - 2A + 2A^2] \\
&= \text{Var}[\xi](1 - 2E_A[A] + 2E_A[A^2]) \\
&= \frac{2}{3}\text{Var}[\xi].
\end{aligned} \tag{A.5}$$

It can be seen that the quadratic expression  $1 - 2E[A] + 2E[A^2]$  introduces the shrinkage factor in the variance. To correct this, the following problem can be a random number  $B \sim F$  such that  $1 - 2E[B] + 2E[B^2] = 1$ .

The approach in this study assumes that  $F = U(0, b)$  where  $b$  is to be determined. The expression then evaluates to

$$1 - b + \frac{2}{3}b^2 = 1. \tag{A.6}$$

The roots of the abovementioned equation is  $b = \{0; (3/2)\}$ . Naturally, we pick the solution  $b = (3/2)$ . In order to implement this change inside the SMOTE algorithm (see [15] for the details), the factor  $\alpha$  can be replaced with  $\tilde{\alpha} = b\alpha$ .  $\square$

## B. Related Figures

The effect of data augmentation from the available original minority of 5 focal ECG intervals in 8 equal steps is shown in Figures 5–7 where 5 ms is taken as the bin size of the histogram.

## Data Availability

The data used to support the findings of the study can be obtained from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

The first and second authors would like to thank Dr. Decebal Gabriel La, tcu from Centre Hospitalier Princesse Grace, Monaco, for contributing the ECG data used for this study.

## References

- [1] J. Granada, W. Uribe, P.-H. Chyou et al., “Incidence and predictors of atrial flutter in the general population,” *Journal of the American College of Cardiology*, vol. 36, no. 7, pp. 2242–2246, 2000.
- [2] F. G. Cosío, A. Pastor, A. Núñez, A. P. Magalhaes, and P. Awamleh, “Atrial flutter: an update,” *Revista Espanola de Cardiologia*, vol. 59, no. 8, pp. 816–831, 2006.
- [3] N. Saoudi, F. Cosio, A. e.-a. Waldo et al., “A classification of atrial flutter and regular atrial tachycardia according to electrophysiological mechanisms and anatomical bases. A statement from a joint expert group from the working group of arrhythmias of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology,” *European Heart Journal*, vol. 22, no. 14, pp. 1162–1182, 2001.
- [4] S. Kanwal, F. Tao, A. Almogren, A. Ur Rehman, R. Taj, and A. Radwan, “A robust data hiding reversible technique for improving the security in e-health care system,” *Computer Modeling in Engineering & Sciences*, vol. 134, no. 1, pp. 201–219, 2022.
- [5] C. Pedrinazzi, O. Durin, G. Mascioli et al., “Atrial flutter: from ECG to electroanatomical 3D mapping,” *Heart International*, vol. 2, no. 3-4, p. 161, 2006.
- [6] U. R. Acharya, H. Fujita, M. Adam et al., “Automated characterization of arrhythmias using nonlinear features from tachycardia ECG beats,” in *Proceedings of the 2016 IEEE International Conference on Systems, Man and Cybernetics (SMC)*, pp. 533–538, IEEE, Budapest, Hungary, October 2016.
- [7] C. Medi and J. M. Kalman, “Prediction of the atrial flutter circuit location from the surface electrocardiogram,” *Europace*, vol. 10, no. 7, pp. 786–796, 2008.
- [8] J. P. Brown, D. E. Krummen, G. K. Feld, and S. M. Narayan, “Using electrocardiographic activation time and diastolic intervals to separate focal from macro-re-entrant atrial tachycardias,” *Journal of the American College of Cardiology*, vol. 49, no. 19, pp. 1965–1973, 2007.
- [9] J. Lee and S. P. Fynn, “P wave morphology in guiding the ablation strategy of focal atrial tachycardias and atrial flutter,” *Current Cardiology Reviews*, vol. 11, no. 2, pp. 103–110, 2014.
- [10] S.-L. Chang, H.-M. Tsao, Y.-J. Lin et al., “Differentiating macroreentrant from focal atrial tachycardias occurred after circumferential pulmonary vein isolation,” *Journal of Cardiovascular Electrophysiology*, vol. 22, no. 7, pp. 748–755, 2011.
- [11] G. Luongo, S. Schuler, A. Luik et al., “Noninvasive characterization of atrial flutter mechanisms using recurrence quantification analysis on the ECG: a computational study,” *IEEE Transactions on Biomedical Engineering*, vol. 67, no. 5, pp. 1234–1243, 2020.
- [12] G. Luongo, *Atrial arrhythmia diagnosis using the 12-lead ECG: Machine learning leveraging in silico and clinical signals*, Ph.D. Dissertation, Karlsruhe Institut für Technologie (KIT), Karlsruhe, Germany, 2021.
- [13] G. M. Weiss, “Mining with rarity: a unifying framework,” *ACM SIGKDD Explorations Newsletter*, vol. 6, no. 1, pp. 7–19, 2004.
- [14] R. C. Holte, L. Acker, and B. W. Porter, “Concept learning and the problem of small disjuncts,” in *Proceedings of the International Joint Conference on Artificial Intelligence*, vol. 89, pp. 813–818, Vienna, Austria, August 1989.
- [15] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, “SMOTE: synthetic minority over-sampling technique,” *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, 2002.
- [16] L. Sun, Z. Shang, Q. Cao, K. Chen, and J. Li, “Electrocardiogram diagnosis based on SMOTE+ENN and random forest,” in *Proceedings of the International Conference on Intelligent Computing*, pp. 747–757, Shandong, China, December 2019.
- [17] R. Blagus and L. Lusa, “SMOTE for high-dimensional class-imbalanced data,” *BMC Bioinformatics*, vol. 14, no. 1, p. 106, 2013.

- [18] S.-S. Bun, D. G. Latcu, F. Marchlinski, and N. Saoudi, "Atrial flutter: more than just one of a kind," *European Heart Journal*, vol. 36, no. 35, pp. 2356–2363, 2015.
- [19] M. H. K. Azman, O. Meste, and K. Kadir, "Detecting flutter waves in the electrocardiogram using generalized likelihood ratio test," vol. 45, pp. 1–4, in *Proceedings of the 2018 Computing in Cardiology Conference (CinC)*, vol. 45, pp. 1–4, IEEE, Maastricht, Netherlands, September 2018.
- [20] M. H. K. Azman, O. Meste, K. Kadir, and D. G. Latcu, "Estimation and removal of t wave component in atrial flutter ECG to aid noninvasive localization of ectopic source," in *Proceedings of the 2017 Computing in Cardiology (CinC)*, pp. 1–4, IEEE, Rennes, France, September 2017.
- [21] C. Seiffert, T. M. Khoshgoftaar, J. Van Hulse, and A. Napolitano, "Rusboost: improving classification performance when training data is skewed," in *Proceedings of the 2008 19th International Conference on Pattern Recognition*, pp. 1–4, IEEE, Tampa, FL, USA, December 2008.
- [22] A. Fernández, S. Garcia, F. Herrera, and N. V. Chawla, "SMOTE for learning from imbalanced data: progress and challenges, marking the 15-year anniversary," *Journal of Artificial Intelligence Research*, vol. 61, pp. 863–905, 2018.
- [23] B. W. Silverman and G. A. Young, "The bootstrap: to smooth or not to smooth?" *Biometrika*, vol. 74, no. 3, pp. 469–479, 1987.
- [24] B. Efron, "Bootstrap methods: another look at the jackknife," in *Breakthroughs in Statistics*, pp. 569–593, Springer, Berlin, Germany, 1992.



## Research Article

# A Privacy Protection Framework for Medical Image Security without Key Dependency Based on Visual Cryptography and Trusted Computing

Denghui Zhang,<sup>1,2</sup> Lijing Ren,<sup>2</sup> Muhammad Shafiq ,<sup>1</sup> and Zhaoquan Gu<sup>2,3</sup>

<sup>1</sup>Cyberspace Institute of Advanced Technology, Guangzhou University, Guangzhou 510000, China

<sup>2</sup>Department of New Networks, Peng Cheng Laboratory, Shenzhen 518055, China

<sup>3</sup>School of Computer Science and Technology, Harbin Institute of Technology (Shenzhen), Shenzhen 518055, China

Correspondence should be addressed to Muhammad Shafiq; srsshafiq@gmail.com

Received 9 August 2022; Revised 30 September 2022; Accepted 24 November 2022; Published 31 January 2023

Academic Editor: Inam Ullah

Copyright © 2023 Denghui Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The development of mobile Internet and the popularization of intelligent sensor devices greatly facilitate the generation and transmission of massive multimedia data including medical images and pathological models on the open network. The popularity of artificial intelligence (AI) technologies has greatly improved the efficiency of medical image recognition and diagnosis. However, it also poses new challenges to the security and privacy of medical data. The leakage of medical images related to users' privacy is emerging one after another. The existing privacy protection methods based on cryptography or watermarking often bring a burden to image transmission. In this paper, we propose a privacy-preserving recognition network for medical images (called MPVCNet) to solve these problems. MPVCNet uses visual cryptography (VC) to transmit images by sharing. Benefiting from the secret-sharing characteristics of VC, MPVCNet can securely transmit images in clear text, which can both protect privacy and mitigate performance loss. Aiming at the problem that VC is easy to forge, we combine trusted computing environments (TEE) and blind watermarking technologies to embed verification information into sharing images. We further leverage the transfer learning technology to abate the side effect resulting from the use of visual cryptography. The results of the experiment show that our approach can maintain the trustworthiness and recognition performance of the recognition networks while protecting the privacy of medical images.

## 1. Introduction

Recently, smart medicine has become an attractive field of applications with the development of 5G, IoT, and AI [1]. Telemedicine has come a long way in detecting and diagnosing diseases remotely, which means that medical images are transmitted more frequently over the open network. The gradual combination of modern medicine with computer technology, communication technology, and multimedia technology has provided patients and doctors with fast medical diagnoses. These technologies have greatly improved the accuracy of medical diagnosis and relieved patients who enjoy the convenience of digital medicine. The storage and analysis of medical images are gradually moving

to the cloud, which is a prerequisite for efficient cooperation in remote diagnosis and resource sharing [2].

Medical imaging devices facilitate doctors' diagnosis and treatment. Benefiting from advances such as 5G in wireless communications and IoT in the industrial Internet, it is now easier to capture and transfer images in the medical field [3]. However, these novel technologies pose new challenges to patients' privacy [4]. Medical images are usually characterized by huge data volume and high security, which contain a large amount of personal information when they are not desensitized. For example, German security firm Greenbone Networks discovered 24 million leaked medical images in 2020, and two months later, the number of exposed scans exceeded 1.19 billion. Frequent medical data

leakage incidents have seriously violated users' privacy and endangered social security. Therefore, privacy protection has to become a requisite for e-healthcare systems [5]. Although different types of data protection algorithms have been proposed, most of them are aimed at text or digital media, which is different from medical images in information quantity and scale [6]. It is often unnecessary to encrypt all pixels to securely transmit an image [7].

Traditional encryption methods based on public keys not only need complex computation but also need additional key management [8]. A common obstacle to these methods is that once the key is lost, we cannot restore the secret of encrypted information [9]. Edge image capture devices generally have limited computing power, which cannot meet the demand for real-time encryption of large volumes of images. These shortcomings limit the spread of AI technology in healthcare. Strengthening the security and protection of privacy of medical images without performance degeneration is urgent. Big data is the key to the success of medical AI. The proposal of a lightweight image protection scheme is the main challenge in this article [10].

Visual cryptography (VC) [11, 12], which is a secret-sharing technology aimed at images, can be applied to digital devices with limited computing in an untrusted networking environment. Using the threshold and secret-sharing features of VC, we can use simple Boolean operations to achieve real-time encryption of large amounts of image data while removing the dependence on keys. However, the VC-generated share is not easy to manage and is vulnerable to attackers.

To preserve the privacy of data when sharing medical data, we present the MPVCNet (medical privacy-oriented VC-based recognition network) to address illegal access and identity forgery for sharing of medical images. We first construct a verifiable and expansion-free visual cryptography scheme to migrate easily forged and size-expanded shares by combining the VC and TEE technologies, where TEE is used to ensure that remote sharing operation is not tampered with [13]. Then, we can securely store and transmit medical images distributed with the proposed scheme. Since the separated sheets do not reveal any feature of the original biomedical data, we can transmit sensitive images among public networks in plain view. MPVCNet eliminates complex computing operations and key management workload in the traditional public key or watermark protection method through the software and hardware cooperation scheme.

## 2. Related Work

VC is a secret-sharing mechanism for images. Since it was first proposed in 1995, it has become emerging research in the field of image security [11, 14, 15]. Unlike the text-oriented secret-sharing mechanism, which divides the sum into two addends, VC splits secret pixels into multiple subpixels. All black pixels will remain once overlapping sharing blocks, although degrading white secret pixels, the

difference of contrast between black and white pixels makes human eyes still recognize the features of the restored image [16]. Because a single pixel is encrypted into a larger color block, the size of the encrypted and decrypted image will expand. The noise-like shares are also often exploited by malicious users.

Because of the low-pass filtering characteristic of human eyes, an image can be recognized even if pixels in the local area change. The probabilistic VCS uses a pixel instead of a color block to encrypt the secret pixel with this feature, thus keeping the size of the images consistent [17, 18]. Although the wrong pixel may be selected at one time, the probability ensures that when many pixels are gathered, the displayed image has the same probability density as the original image; that is, the two images are similar. If we encrypt one color block at a time instead of one pixel, we can also keep the size of the decrypted image inconvenient [19]. By formulating the correspondence between secret blocks and color blocks before and after encryption, we can decrypt higher-quality secret images [10]. However, this mechanism often leads to more shares, which requires more storage space. The software-based encryption scheme is hard to avoid the attack on the hypervisor and remote operating system.

Since images generated by VC are meaningless, the attacker can forge shares without being detected. To address this limitation, researchers have proposed many CIVCS (cheating immune VCS) [20–22]. If there is no checksum information, malicious users can forge shares and deceive users. So, CIVCS often requires additional pixels or more shares to embed verification information, which can burden the cost of VC.

Big data and cloud computing technologies have solved two main challenges in the spread of medical AI [23, 24]. A growing number of smart medical systems facilitate people's daily life [25], which has become an important research direction in healthcare. The authors [26] use deep learning algorithms to identify diabetic retinopathy lesions, which are the most common sequelae of diabetes and can lead to blindness. The experimental results show that the AI algorithm outperforms medical experts and can extract lesion features from fine textures. Due to the difficulty and scarcity in the acquisition of medical image annotation data, transfer learning [27] has become a very common technique for medical image recognition. TransFusion [25] investigates the evolution of representations of different models and hidden layers during training and the advantage of feature independence of migration learning to accelerate convergence.

## 3. MPVCNet

With the introduction of data security regulations and increasing awareness of the limitations of AI technology, it is imperative to ensure the privacy of personal information when enjoying the convenience of AI technology [4]. To address these limitations, we will first present a size-invariant and verifiable VCS and then propose a secure and trustworthy image recognition network in this section.

3.1. *The List of Abbreviations.* Table 1 provides the list of abbreviations used in this article.

### 3.2. Backgrounds

3.2.1. *Visual Cryptography Schemes (VCSs).* The secret-sharing scheme, which is also called a threshold scheme, was originally used to provide a solution when many participants needed to share a secret. Only the number of participants who reach a threshold can recover sensitive information such as pathological and CT images. Participants less than the threshold cannot reveal any information about the original image even if they conspire.

As an extension of secret-sharing technology toward image encryption, the principle of VC is to divide an image into  $n$  unrelated shares (also called sheets), which are independent of each other and distributed to  $n$  individuals for safekeeping. If  $m, m \leq n$ , image cannot reconstruct the original image, thus achieving image security. Image encryption schemes based on secret-sharing have the advantages of high security and simple computation.

VC essentially makes use of the contrast characteristics of human eyes to color. For white pixels, the subpixels of shares 1 and 2 are the same, while for black pixels, the subpixels of shares 1 and 2 are complementary. When superimposing, all-black decrypted pixels can be obtained, while the white decrypted pixel contains two black pixels, but its brightness is still higher than that of the black decrypted pixel.

Equations (1) and (2) illustrate the encryption and decryption process of a (2, 2)-VCS. When splitting a secret image into two meaningless shares, the original image can only be restored if two shares are obtained simultaneously. The  $C_k, k \in \{0, 1\}$  is the encryption matrix for a black pixel (0) and white pixel (1), which can be generated by rotating the base matrix  $S_k$ . The top and bottom rows of the black matrix ( $C_0$ ) can be superimposed to obtain 4 black pixels, and the white matrix ( $C_1$ ) can be superimposed to obtain 2 black pixels. So, the decrypted image looks darker. Therefore, we can visually experience a significant difference at grayscale levels when sharing images using the overlay.

$$S^0 = \begin{bmatrix} [0 & 1] \\ [0 & 1] \\ [0 & 1] \\ [0 & 1] \end{bmatrix}, S^1 = \begin{bmatrix} [0 & 1] \\ [0 & 1] \\ [1 & 0] \\ [1 & 0] \end{bmatrix}, \quad (1)$$

$$C_0 = \{\text{permutation of } S^0\}, C_1 = \{\text{permutation of } S^1\}. \quad (2)$$

3.2.2. *Intel SGX.* Intel Software Guard Extensions (SGX) is one of the most popular TEE technologies [28]. It is an instruction extension proposed by Intel to provide hardware-level protection for user code and data based on the Intel CPU architecture. SGX is a specific implementation of TEE to prevent other applications, including privileged operating systems, from tampering with the user code and sniffing information about applications running in a

TABLE 1: The list of abbreviations used in this article.

Abbreviations	Terms
MPVCNet	Medical privacy-oriented VC-based recognition network
VC	Visual cryptography
IoT	Internet of things
VCS	VC scheme
CIVCS	Cheating immune VCS
SGX	Software guard extensions
SSIM	Structure SIMilarity
PSNR	Peak signal-to-noise ratio
ROC	Receiver operating curve

protected environment (called enclave). Neither privileged nor unprivileged software can access the enclave. SGX technology consists of two core mechanisms: isolated execution and remote authentication.

(1) *Isolated Execution.* An enclave can protect the confidentiality and integrity of the processes running in it, which is defended from attacks by other processes on the host, and blocks hardware-oriented attacks. As shown in Figure 1, SGX uses an inverted sandbox model to protect applications and data from being skewed. Instead of identifying and isolating all malwares on the platform, SGX protects legitimate software by encapsulating its security operations in an enclave and protecting it from malware attacks.

SGX provides only a limited amount of memory (128 MB), and excluding the driver, the space available to the user enclave is even less, equal to about 93 MB [29]. Therefore, at runtime, SGX will load data into the enclave from the CPU to memory as needed or export to memory in ciphertext [30]. The key is stored in hardware and is not accessible even by the operating system. The malicious program can also only sniff the encrypted ciphertext data from memory. When the user program needs to access the previous data, the data in memory are paged back to enclave and the decryption operation is performed. The whole process is transparent to the user. Because the whole encryption and decryption process is performed in hardware, the impact on program performance is small.

(2) *Remote Authentication (RA).* Intel SGX provides a remote authentication mechanism for applications running on remote or cloud platforms. Using the remote authentication mechanism, before the client encrypts and transmits sensitive data to the application running in an enclave, the client will verify the credibility of the remote environment and application. The enclave can indicate to the challenger that its identity has not been altered while running in a trusted Intel SGX environment [31].

In a complete authentication process, a server runs in a trusted SGX environment. At first, the client challenges the server to prove its identity. If an application wants to transmit data to an enclave or receive the execution result of the enclave in confidence, it requires the establishment of a secure transmission channel between an external application and the enclave. The principle is like TLS (transport layer security) of HTTPS in a browser. First, the client and server

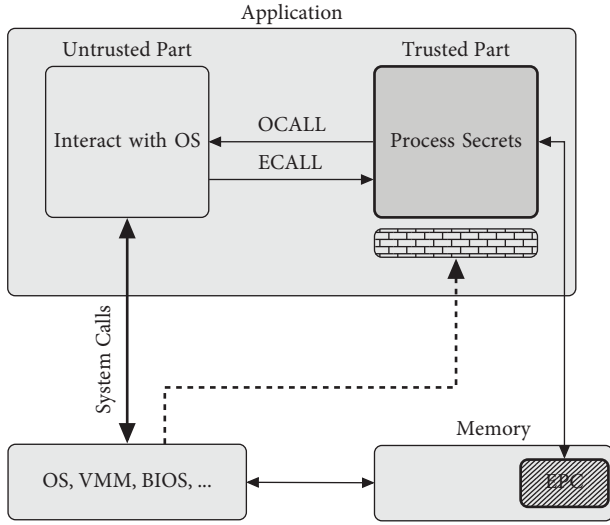


FIGURE 1: The workflow of Intel SGX.

use the key exchange algorithm such as the Diffie–Hellmann key exchange to negotiate the key of the asymmetric encryption algorithm, and then both communication parties use this key to encrypt messages for communication. After passing the Intel attestation service (IAS) verification, the client provides sensitive data to the server in digital copyright protection or approves the operation performed by the server by using the encryption channel established in the authentication process [32]. Please refer to documents [33] for more details about Intel SGX.

**3.2.3. Verifiable Visual Cryptography.** We chose the random grid (RG) method [15] as the backbone scheme; that is,  $C_0$  and  $C_1$  are used as the underlying base matrix. To distribute each column of collections to a participant, we shuffle the subpixel in a column and distribute one of them to a participant instead of the whole column. Although there are more complex VCS [34], which often require more shares and can reduce the quality of the restored image, the advantage of VCS is that it distributes subpixels to corresponding participants; that is, the number of subpixels received by participants is the same as that of encrypted secret pixels, so VCS can keep the image size. While the origin VCS distributes subpixels to one of the participants, it substantially expands the size of the shared image. The encryption is shown in Algorithm 1, which is explained in more detail as follows:

(1) *Decomposition.* Like other color images, medical images often have three (red, green, and blue) gray-level channels. The RGB mode is additive color, which is based on the superposition of light. The red light plus green light and blue light are equal to the white light in nature. This color mode is suited to appropriate devices like displays. While printers utilize ink to absorb light to display a specific color, this color mode is called the subtractive mode, which also has three (cyan, magenta, and yellow) channels. The cyan pigment blends magenta and yellow pigments to obtain black (gray) in printing. The digital

color images are stored in the RGB mode. We first need to decompose the original image into three CMY gray-level channel images. This can be archived from RGB channels with the following transformation to get complementary colors:  $C = 255 \setminus R$ ,  $M = 255 \setminus G$ ,  $Y = 255 \setminus B$ .

(2) *Digital Half-Toning.* The original VCS only can encrypt white and black pixels. While the gray-level channel images have 256 levels (as a reference, the binary image has only two gray levels). Halftone is a commonly-used digital image processing technology, which simulates continuous tone images by differing the frequency of a small amount of color, and the visual effect of the quantized image is like the original image at a certain distance. Halftone technology is based on the low-pass filtering feature of human eyes and color characteristics to achieve optimal image reproduction in monochrome or binary color equipment. When viewing spatially close parts of an image, eyes will form the effect of continuous tone in general. Error-diffusion [35] is a commonly used halftone algorithm that can be described as follows:

$$\begin{cases} p_{x-1,y+1} = p_{x-1,y+1} + qe \times c[1,0], \\ p_{x+1,y} = p_{x+1,y} + qe \times c[0,2], \\ p_{x+1,y} = p_{x,y+1} + qe \times c[1,1], \\ p_{x+1,y} = p_{x+1,y+1} + qe \times c[1,2], \end{cases} \quad (3)$$

$$qe = p_{x,y} - T_{x,y}, \quad (4)$$

$$\begin{cases} T_{x,y} = 0, & \text{if } p_{x,y} < \tau, \\ T_{x,y} = 1, & \text{if } p_{x,y} \geq \tau, \end{cases} \quad (5)$$

$$c = \begin{bmatrix} p_{\text{old}} & p & \frac{7}{16} \\ \frac{3}{16} & \frac{5}{16} & \frac{1}{16} \end{bmatrix},$$

where  $qe$  is the optimal quantization error and it usually is a threshold function truncated by  $\tau$ . The coefficients  $c$  are optimal arrays to make diffusion smoother, which are  $[7/16, 3/16, 5/16, 1/16]$ . Considering the gray value of the neighborhood pixels, the output image is smoother and the noise is less than that of the random or jitter method.

(3) *Encryption.* The encryption process is applied to each halftone channel image separately. When processing a pixel  $p_{ij}$ , we will set the corresponding pixel  $S_{1ij}$  in *share1* to be a white or black pixel randomly. If  $p_{ij}$  is a white pixel, we set the corresponding pixel  $S_{2ij}$  in *share2* to be the same as  $S_{1ij}$ . While if  $p_{ij}$  is a black pixel, we will set  $S_{2ij}$  to be the inverse of  $S_{1ij}$ . For example, for the black, we will reinterpret  $S_{2ij}$  as a white pixel if  $S_{1ij}$  is also a black pixel. While if  $S_{1ij}$  is a black pixel,  $S_{2ij}$  will be a white pixel. After generating and sharing images  $S_1$  and  $S_2$ , we will generate a corresponding verification message. First, a random number  $R$  is generated so that the receiver can use  $R$  to perform pair operations when receiving multiple shares. The user can also use a GUID

```

Data: A secret image  $S$ 
Result: Share images  $S_1, S_2$  dispatched to participants  $P_1, P_2$ 
 $S_i \leftarrow$  the half-toned secret image got with the error-diffusion algorithm from equation (1)
for each pixel  $p_{ij}$  in  $S_i$  do
  if  $p$  is black then
     $d_1, d_2, \dots, d_n \leftarrow n$  randomly generated black or white pixels
     $S_{nij} = d_n, n$  is the number of shares
  end
  else
     $d \leftarrow$  a randomly generated black or white pixel
     $S_{nij} = d$ 
  end
end
 $R \leftarrow$  a random number
for each share  $S_i$  do
   $A \leftarrow S_i \parallel R$ 
   $Q \leftarrow$  a generated quote that signs the hash value of  $A$ 
   $AR \leftarrow Q \parallel R$ 
for each bit  $b_j$  in  $AR$  do
   $p_j \leftarrow$  the corresponding pixel in  $S_i$ 
   $\text{LSB}(p_j) \leftarrow b_j$ 
end
end

```

ALGORITHM 1: A verifiable visual cryptography scheme.

(globally unique identifier) to indicate pairing if it can be uniquely identified in this system. The enclave ( $E$ ), which is the trusted part of the encryption program, and *quote* enclave ( $Q$ ) in the host will first perform local authentication.  $E$  asks  $Q$  to sign the hash value of the spliced bitstream  $A = S_i \parallel R$ , and we embed  $A$  into the quote structure. Since the hash value is fixed, the length of the final generated quote structure is also fixed. Finally, the authentication code  $AR = Q \parallel R$  is generated, and each bit in  $AR$  is assigned to LSB (lowest significant bit) for each pixel, in turn, starting from the first pixel of the corresponding share. For example, if  $AR[i] = 1$  and  $S_1[i]$  are white pixels, then we will set  $\text{LSB}(S_1[i])$  to 1, the grayscale of the shared pixel  $S_1[i]$  will be reassigned to 255; that is, it is still a white pixel. While if  $AR[i] = 0$ ,  $S_1[i]$  will be reassigned to 254 which is received by setting the lowest bit of 255 to 0. When the pixel in  $S_1$  is black, the grayscale value of the corresponding pixel becomes 0 or 1. So, the range of pixels in shares after embedding an authentication code value is [0, 255, 1, and 254]. The embedding process of the authentication information into the share  $S_2$  is similar.

(4) *Check and Decryption.* When restoring the original image, we can ask participants  $P_1$  and  $P_2$  to show their sheets. Before overlaying shares to recover the secret image, we first need to extract the verification information from the received shares. Since the *quote* structure and the paired random number ( $R$ ) are both of fixed length, the embedding information is hidden in the first  $n$  (the sum of the byte lengths of  $Q$  and  $R$ ) bits of the received shares. The extraction of verification information is an inverse process of the previous embedding process. Taking  $S_1$ , for example, if the  $i$ th bit of  $S_1$  is 1 or 255, the  $i$ th bit will be 1 in the

corresponding verification information ( $AR$ ). While if  $S_1$  is 0 or 254, we will assign  $AR[i]$  to 0. Once the extraction process is completed, we will get the quote and pairing information and recover the pixels with a grayscale of 1 or 254 in shares to 0 and 255, respectively. We take received images as trusted shares only if they satisfy the following conditions: (i) the pairing information of the shares is the same; (ii) the validity of  $Q$  is requested to be verified by IAS, and the validity of the IAS report is verified by the local certificate. If the report is valid and the signature information of this report is consistent with the locally stored IAS certificate information and the *MRENCLAVE* ID extracted from  $Q$  is in the list of trusted enclave IDs, then we think  $Q$  is trusted; (iii) the hash value of the joined bitstream of  $Q$  and  $R$  is consistent with the embedded hash information in  $Q$ . After passing all of the previous checks, we can print out the two shares and superimpose them together or simply perform a Boolean *AND* operation to reveal secret images. Note that we perform Boolean *AND* while the Boolean *OR* operation is executed in the original scheme. This is because the original scheme uses 1 for black pixels and 0 for white pixels, whereas white pixels are generally represented by 1 or 255 and black pixels by 0 in normal image files.

3.2.4. *The Recognition Network of Medical Images.* In the preceding section, VC provides a safe and effective encryption scheme for the transmission and analysis of large-scale medical image data sets, which eliminates the key dependence on traditional encryption methods. Although the application of VC can efficiently transmit medical images, the quality of restored images may be lossy. This may break recognition performance [36]. In this section, we will

use the strong feature extraction ability of deep learning to propose a high-accuracy recognition network of noisy or blurred medical images.

Traditional image recognition methods obtain recognition models through elaborate feature extraction methods, which are limited by environments such as lighting, contrast, and pose, although human eyes can accurately recognize friends around us, regardless of the external conditions. With the construction of large-scale datasets and increasing computing power, deep learning-based image recognition methods are now capable of recognizing complex scenes with high performance even beyond human eyes [37].

Deep learning-based models can be automatically trained with datasets to extract complex mapping relationships between input images and target labels. However, in the process of image recognition, to maximize the recognition performance of neural networks, it is generally necessary to first design the network structure and loss function according to the features of datasets and then continuously train on the cloud platform until a stable network model is obtained [38, 39]. This process is often time-consuming, and recognition performance is still influenced by the selection of hyperparameters in the training process.

With the further development of machine learning methods, we can now directly reuse the model parameters learned from large-scale datasets. Note that our lossy medical image dataset does not have the same feature distribution as generic datasets such as ImageNet, and it is still not feasible to reuse existing models and weights directly. To issue a problem, we can add a fully connected layer at the end of the network and then fine-tune the model to match the size and class of the medical image data. Then, we can use the method in the previous section to encrypt and decrypt the original training dataset one by one to obtain a recognition model for the lossy dataset. Since this approach considers both the features and weights learned from a large-scale generic dataset and the characteristics of the model itself, therefore we can efficiently obtain a high-accuracy recognition network of noisy or blurred medical images. We only need to add a new classifier to the pretraining model, which will be trained from scratch [40].

Previous experiments [25] show that feature reuse mainly occurs at the lowest level. We conducted a weighted transfusion operation to speed up the training process; that is, only a part of the pretraining weights (corresponding to a group of consecutive layers) was transferred, and the rest weights were randomly initialized. Compared with that of complete transfer learning, the convergence speed of the indoctrination network will further accelerate.

## 4. Results and Discussion

To assess the effectiveness and security of the method proposed in this paper, we will test the performance of the proposed framework in signing encryption and images on the prototype system developed using the C++ language based on the SGX SDK. We evaluate the latency details of the encrypted and signed multiple images by using a modified

SGX-SSL benchmarking tool. We performed experiments on an Intel desktop CPU with 16 GB RAM running on the Linux system. We both test the performance of the verifiable VCS on a normal host and a host that supports SGX features.

Figure 2 shows that the normal operation had a higher performance where no security enhancements are adopted. There are two reasons why the performance of signing in TEE is weaker than the plain signing: (i) one is the loss due to the use of encryption memory in TEE, which is inevitable in order to prevent the privileged system from sniffing; (ii) the other is the extra function calls resulting from the refactoring of the original code to allow the original code to run in TEE. However, the loss from direct function calls is still less than the loss from *ECALL/OCALL* in the enclave running on a host that allows the SGX feature.

Figure 3 shows the latency between the TEE-based approach presented in this paper and other commonly used image encryption methods. Although SGX has a higher performance payload than conventional transmission, the loss is lower than the encryption methods, including RSA and TLS. The more complex the encryption process, the higher the efficiency of encryption, which further shows the importance of adopting visual cryptography and the TEE software and hardware coprotection method. The performance difference can be attributed to two reasons. (i) SGX is an extended instruction set integrated with the CPU that can call hardware instructions to speed up the encryption and decryption process. (ii) SGX distinguishes between trusted and untrusted codes; not only is the normal operation split into two sets of code, but the parameters must be toggled and copied back and forth between the two running spaces, which enhances security at an extra cost. So, encryption, decryption, and signing operations with SGX are more time-consuming than in a normal environment, but the average loss of  $8\times$  is within reasonable bounds.

We further develop an application using the TensorFlow [41] framework to test the recognition network model rapidly. We first select the diabetic retinopathy (DR) detection dataset to identify signs of diabetic retinopathy in eye images, which consists of 35126 DR images with the size  $786 \times 512$ . The DR is the first cause of blinding. We choose another *BreakHis* [42] (breast cancer histopathological) dataset which contains 7909 breast histopathological images from 82 patients to further evaluate the method proposed. *BreakHis* not only advances the research of binary classification of benign and malignant but also advances the research of pathological classification (multiclassification), which is very significant in the clinic. Since the size of the images in the dataset varies, the shape of the input features of the neural network needs to be consistent. Therefore, to simplify the experiment, we used the smallest image in the dataset as the benchmark and then used data enhancement to obtain the preprocessed dataset with  $S = \{1, 2, 4, 5\}$ ; of these, 70% were used as a training set, and 30% as a test set shows the data distribution of the DP dataset for different levels. The third and fourth columns of Table 2 are the average *PSNR* (peak signal-to-noise ratio) from half-toned and recovery images in each class according to equation (4). Due to the use of half-toned and color confusion

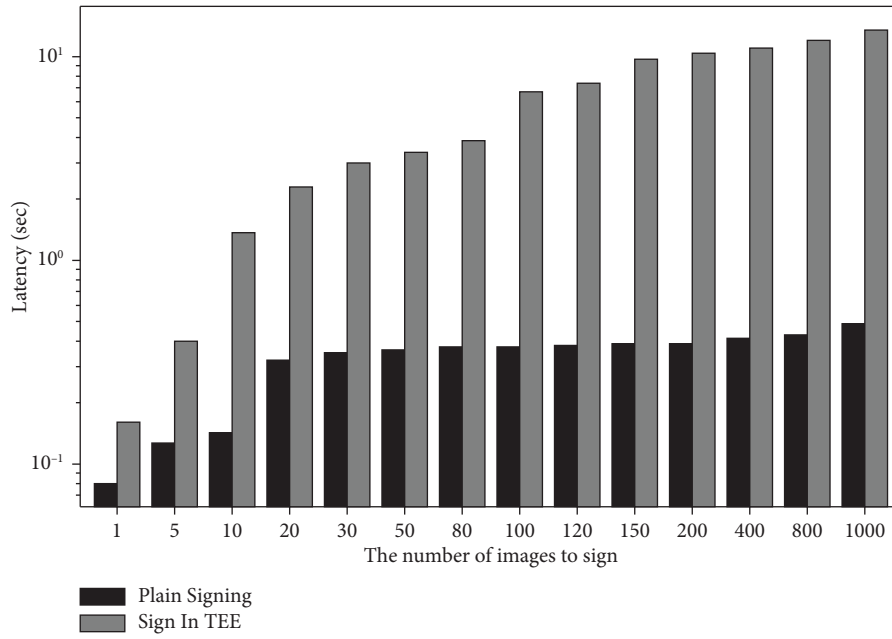


FIGURE 2: Comparison of the signing performance between normal and TEE environments.

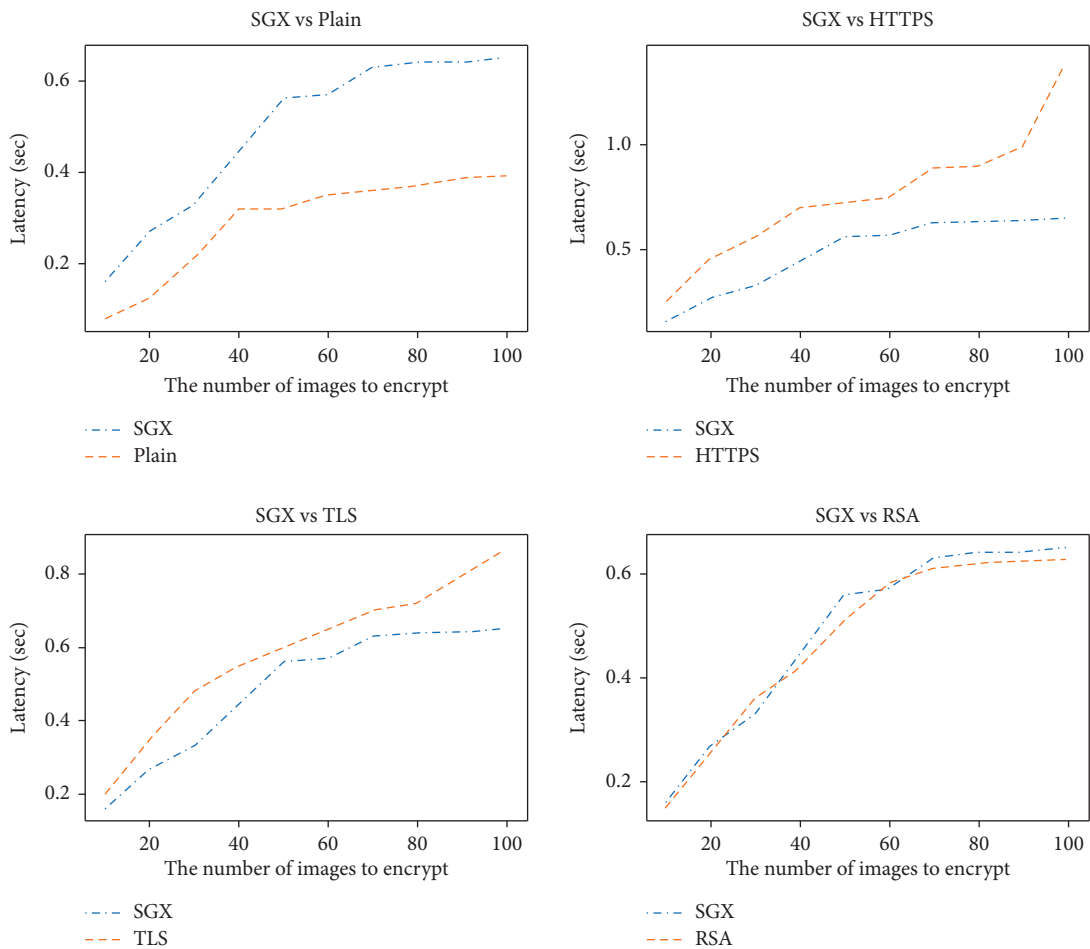
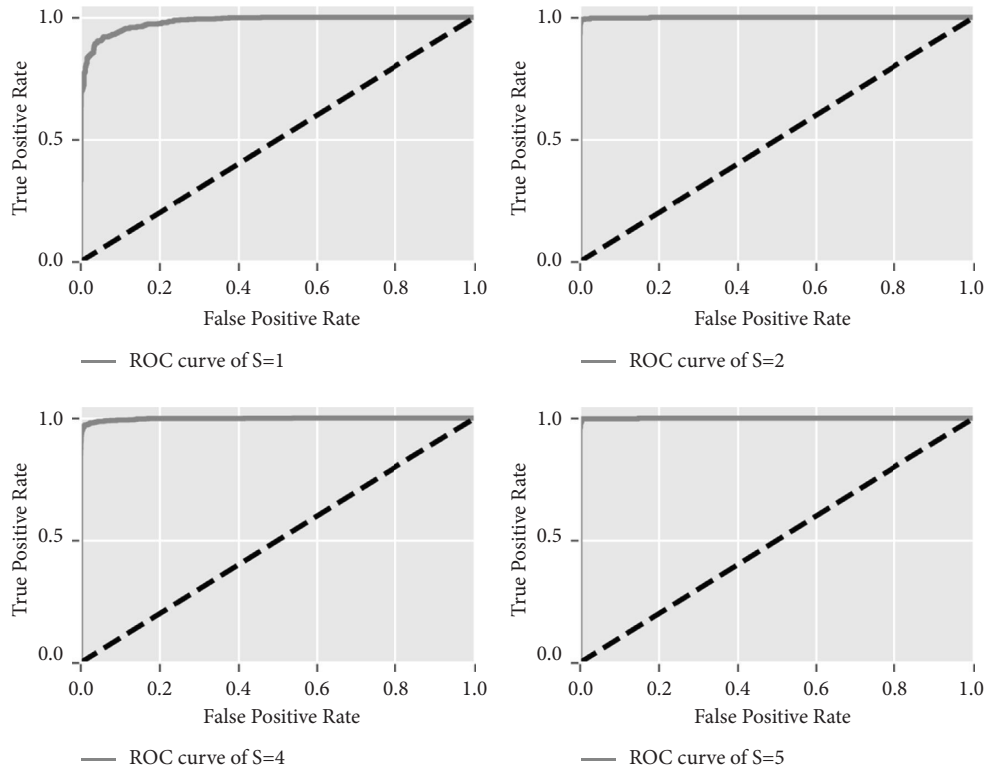


FIGURE 3: Comparison of SGX and other encryption methods.

TABLE 2: Data distribution and image quality of the DP dataset.

Class	Count	Halfone’s PSNR	Our PSNR	Our SSIM	Wu’s SSIM
No DR (0)	25810	27.77	27.77	7.26	2.07
Mild (1)	5292	27.76	27.71	3.34	1.32
Moderate (2)	2443	28.25	28.00	3.88	1.51
Severe (3)	873	28.17	28.13	1.78	0.92
Proliferative DR (4)	708	27.83	27.76	11.08	3.11
Total/average	35126	27.95	27.87	5.47	1.79

FIGURE 4: The ROC curves of the *BreakHis* dataset from normal images.

technologies, the quality of restored images is degraded. The image quality is almost equal to the PSNR quality of the half-toned image. The PSNR values are about 28, which does not affect the visual effect. We chose the SSIM (Structure SIMilarity) metric to further evaluate the decrypted image quality with the proposed method.

As shown in Table 2, the SSIM distinguishes images with global noise such as half-tones better than PSNR. Because our verifiable VCS is based on the random grid method, which is a lossless encryption method, although some image information is lost because of embedding TEE-based authentication information, the decryption quality of our method is still better than Wu’s method [18], and the SSIM metric is about  $3\times$  that of Wu’s method.

We choose another *BreakHis* [42] (breast cancer histopathological) dataset which contains 7909 breast histopathological images from 82 patients to further evaluate the method proposed. *BreakHis* not only advances the research of binary classification of benign and malignant but also advances the research of pathological classification (multi-classification), which is very significant in the clinic. Since

the size of the images in the dataset varies, the shape of the input features of the neural network needs to be consistent. Therefore, to simplify the experiment, we used the smallest image in the dataset as the benchmark and then used data enhancement to obtain the preprocessed dataset with  $S = \{1, 2, 4, 5\}$ , and of these, 70% were used as a training set and 30% as a test set.

In the process of medical image recognition, an important factor is the resolution of the image. For example, the classic *BreakHis* provides four scales of datasets, and the visual effects of encryption and decryption of images with different scales are different. To test the performance of the proposed method in multiscale datasets, we show ROCs with four resolutions in Figures 4 and 5, respectively.

Finally, we compare the recognition performance of the normal dataset with the restored dataset using the *ResNet18* [43] backbone on the *BreakHis* dataset. Figure 4 (normal images) and Figure 5 (recovery images) show the recognition performance toward the *BreakHis* dataset in the ROC curves. The AUCs from normal and recovery datasets are 0.95 and 0.93, respectively. Despite a slight decline in



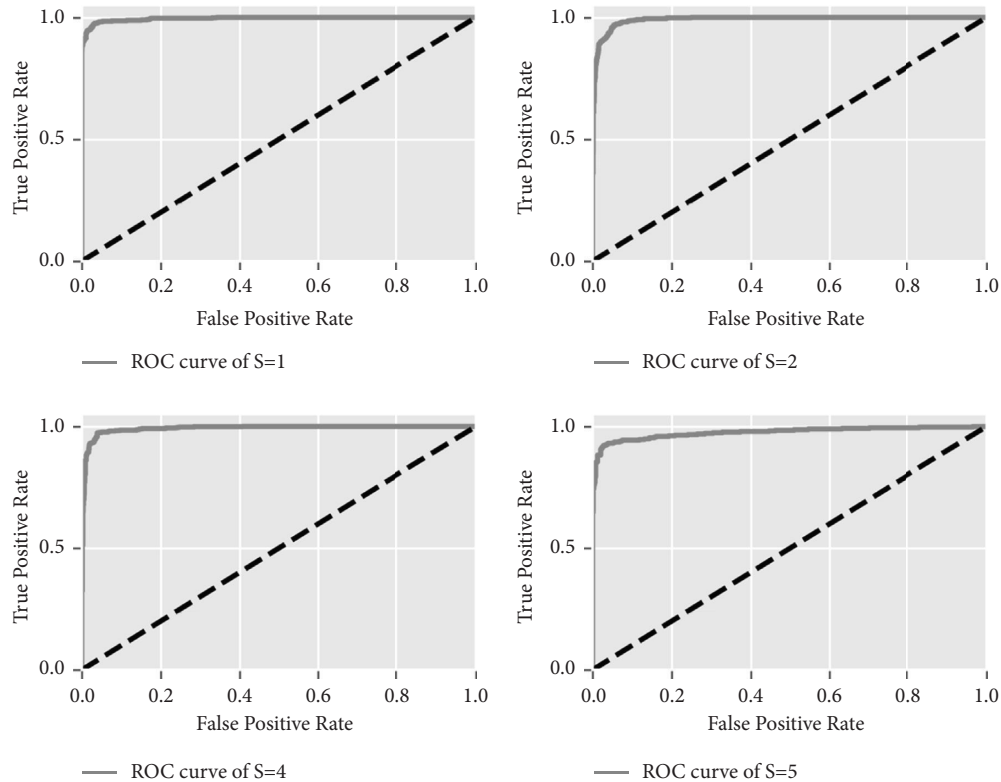


FIGURE 5: The ROC curves of the *BreakHis* dataset from recovery images.

performance, this experiment indicates that our find-tuned can archive high recognition even in the loss medical image dataset.

**4.1. Security Analysis.** As shown in the MPVCNet section, the proposed MPVCNet in this paper is divided into four parts, among which decomposition and digital half-toning belong to preprocessing, which will not affect the security of the encrypted image. In the encryption section, the existing visual cryptographic algorithms are reused, and these algorithms have been proven to be secure. To prevent the shared data from being forged, we use TEE technology to sign the encrypted shares. The whole signing process is executed in a trusted enclave. The signature key can only be accessed by a verified enclave. Even if the operating system is attacked or hijacked, it will only lead to the failure of the signature process and it is impossible to generate forged shares. In the final encryption process, the signature information will be extracted first, and then IAS will be requested for verification. Once the signature is found to not match, the system will terminate immediately. MPVCNet organically combines VC with TEE and digital signature, which ensures the efficiency and security of the whole image transmission process.

The proposed scheme maintains the visual recovery property of the VC. If the user does not have a computing device, we can still recover secret images by printing and stacking shares directly without extracting the validation information. Because we only use the lowest of the shared image to store the authentication information, the final

lost gray value is less than  $1/255$ . The complete certificate chain including local certificates, remote IAS, trusted enclave list, and pairing information during the authentication process can prevent malicious users from tampering with shared information and thus falsifying medical data. Unlike previous TLS or HTTPS-based encryption, this scheme does not require the preestablishment of a trusted channel. The entire transmission can be performed offline except that a request to the IAS is required to verify the validity of the quote. This feature facilitates the transmission of medical images since many of which are transmitted by papers or transparencies. Furthermore, unlike other trusted third-party-based authentication methods, this mechanism uses the TEE feature to remove the reliance on third-party certificate authorities. The locally maintained *MRENCLAVE* list can also be used for revocation. If an enclave explodes with a vulnerability, we can remove the ID of an enclave from the list, and then the shares generated by that enclave will not be trusted.

## 5. Conclusions

To address security issues after the spread of smart healthcare, we explore the possibility of using visual cryptography to preserve the privacy of medical data in this document. Simple VCS provides an effective method to allow distributed storage of images, which avoids centralized storage risks in the cloud environment. VCS is suitable for smart devices with low computing power. While existing VCS for imparting privacy of

biometric templates violates the principle, by encrypting a medical image pixel by pixel and issuing subpixels in one column of selection matrices to the corresponding participants, we archive visual cryptography without expanding the image size and keeping the computation-free feature of VCS. Although there have been various cheating-immune VCSs, we use the characteristics of TEE to effectively ensure that the sharing operation has not been maliciously tampered with, thus fundamentally ensuring the security of sharing images.

In future works, we will investigate reasons why we can keep the recognition performance for encrypted images by fine-tuning a neural network even when there is noise in the decrypted image. This reaffirms the powerful self-learning capability of deep learning and illustrates the non-interpretable machine learning since the machine learning model can maintain high recognition accuracy despite the perceived degradation of quality by the human eye. Second, we are working on class activation mapping to understand and reveal the decision-making process of the neural network on encrypted data. [5].

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (62250410365), the National Key Research and Development Program of China (2019YFB1706003), the Major Key Project of PCL (PCL2022A03), the Guangdong Key R&D Program of China (2019B010136003), the Guangdong Higher Education Innovation Group (2020KCXTD007), the Guangzhou Higher Education Innovation Group (202032854), the Guangdong Province Universities and Colleges Pearl River Scholar Funded Scheme (2019), the Guangdong Basic and Applied Basic Research Foundation of China (2022A1515011542), and the Guangzhou Science and Technology Program of China (202201010606).

## References

- [1] F. Al-Turjman, M. H. Nawaz, and U. D. Ulusar, "Intelligence in the Internet of Medical Things era: a systematic review of current and future trends," *Computer Communications*, vol. 150, pp. 644–660, 2020.
- [2] H. Liang, M. Li, Y. Chen, T. Yang, Z. Xie, and L. Jiang, "Architectural protection of trusted system services for SGX enclaves in cloud computing," *IEEE Transactions on Cloud Computing*, vol. 9, no. 3, pp. 910–922, 2021.
- [3] M. Eichelberg, K. Kleber, and M. Kämmerer, "Cybersecurity in PACS and medical imaging: an overview," *Journal of Digital Imaging*, vol. 33, no. 6, pp. 1527–1542, 2020.
- [4] F. Alsubaei, A. Abuhussein, V. Shandilya, and S. Shiva, "IoMT-SAF: Internet of medical Things security assessment framework," *Internet of Things*, vol. 8, Article ID 100123, 2019.
- [5] Y. Tan, J. Qin, L. Tan, H. Tang, and X. Xiang, "A survey on the new development of medical image security algorithms," in *Cloud Computing and Security*, X. Sun, Z. Pan, and E. Bertino, Eds., vol. 11065, pp. 458–467, Springer International Publishing, Berlin, Germany, 2018.
- [6] M. Taheri, S. Mozaffari, and P. Keshavarzi, "Face authentication in encrypted domain based on correlation filters," *Multimedia Tools and Applications*, vol. 77, no. 13, pp. 17043–17067, 2018.
- [7] T. M. Thanh and K. Tanaka, "An image zero-watermarking algorithm based on the encryption of visual map feature with watermark information," *Multimedia Tools and Applications*, vol. 76, no. 11, pp. 13455–13471, 2017.
- [8] Z. Gu, H. Li, S. Khan et al., "IEPSBP: a cost-efficient image encryption algorithm based on parallel chaotic system for green IoT," *IEEE Transactions on Green Communications and Networking*, vol. 6, no. 1, pp. 89–106, 2022.
- [9] Z. Gu, L. Wang, X. Chen et al., "Epidemic risk assessment by a novel communication station based method," *IEEE Transactions on Network Science and Engineering*, vol. 9, no. 1, pp. 332–344, 2022.
- [10] Y. Nakatsuka, A. Paverd, and G. Tsudik, "PDoT: private DNS-over-TLS with TEE support," in *Proceedings of the 35th Annual Computer Security Applications Conference*, pp. 489–499, San Juan, PR, USA, December 2019.
- [11] M. Naor and A. Shamir, "Visual cryptography," in *Advances in Cryptology — EUROCRYPT'94*, A. De Santis, Ed., vol. 950, pp. 1–12, Springer Berlin Heidelberg, Berlin, Heidelberg, 1995.
- [12] L. Tan, K. Liu, and X. Yan, "Robust visual secret sharing scheme applying to QR code," *Security and Communication Networks*, vol. 2018, pp. 1–12, 2018.
- [13] M. S. Islam, M. S. Ozdayi, L. Khan, and M. Kantarcioglu, "Secure IoT data analytics in cloud via Intel SGX," in *Proceedings of the 2020 IEEE 13th International Conference on Cloud Computing (CLOUD)*, pp. 43–52, Beijing, China, October 2020.
- [14] J. Mohan and D. R. Rajesh, "Enhancing home security through visual cryptography," *Microprocessors and Microsystems*, vol. 80, Article ID 103355, 2021.
- [15] H.-C. Chao and T.-Y. Fan, "XOR-based progressive visual secret sharing using generalized random grids," *Displays*, vol. 49, pp. 6–15, 2017.
- [16] Y. Dong, X. Huang, and G. Ye, "Visually meaningful image encryption scheme based on DWT and schur decomposition," *Security and Communication Networks*, vol. 2021, 16 pages, 2021.
- [17] Y. Ren, F. Liu, T. Guo, R. Feng, and D. Lin, "Cheating prevention visual cryptography scheme using Latin square," *IET Information Security*, vol. 11, no. 4, pp. 211–219, 2017.
- [18] X. Wu and C.-N. Yang, "Probabilistic color visual cryptography schemes for black and white secret images," *Journal of Visual Communication and Image Representation*, vol. 70, Article ID 102793, 2020.
- [19] D. Zhang, H. Zhu, S. Liu, and X. Wei, "HP-VCS: a high-quality and printer-friendly visual cryptography scheme," *Journal of Visual Communication and Image Representation*, vol. 78, pp. 103–186, 2021.
- [20] L. Zhang, X. Dang, L. Feng, and J. Yang, "Efficient secret image sharing scheme with authentication and cheating prevention," *Mathematical Problems in Engineering*, vol. 2021, 11 pages, 2021.

- [21] X. Jia, D. Wang, Q. Chu, and Z. Chen, "An efficient XOR-based verifiable visual cryptographic scheme," *Multimedia Tools and Applications*, vol. 78, no. 7, pp. 8207–8223, 2019.
- [22] G. Selva Mary and S. Manoj Kumar, "A self-verifiable computational visual cryptographic protocol for secure two-dimensional image communication," *Measurement Science and Technology*, vol. 30, no. 12, Article ID 125404, 2019.
- [23] D. Silver, J. Schrittwieser, K. Simonyan et al., "Mastering the game of Go without human knowledge," *Nature*, vol. 550, no. 7676, pp. 354–359, 2017.
- [24] F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: a unified embedding for face recognition and clustering," in *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 815–823, Boston, MA, USA, June 2015.
- [25] M. Raghu, C. Zhang, J. Kleinberg, and S. Bengio, "Transfusion: understanding transfer learning for medical imaging," 2019, <https://arxiv.org/abs/1902.07208>.
- [26] V. Gulshan, L. Peng, M. Coram et al., "Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs," *JAMA*, vol. 316, no. 22, 2402 pages, 2016.
- [27] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345–1359, 2010.
- [28] T. Hunt, Z. Zhu, Y. Xu, S. Peter, and E. Witchel, "Ryoan: a distributed sandbox for untrusted computation on secret data," *ACM Transactions on Computer Systems*, vol. 35, no. 4, pp. 1–32, 2018.
- [29] F. McKeen, I. Alexandrovich, I. Anati, and D. Caspi, "Intel® software guard extensions (Intel® SGX) support for dynamic memory management inside an enclave," in *Proceedings of the Hardware and Architectural Support for Security and Privacy 2016 on - HASP 2016*, pp. 1–9, Seoul, Republic of Korea, June 2016.
- [30] Y. Shen, Y. Chen, K. Chen, H. Tian, and S. Yan, "To isolate, or to share?: that is a question for Intel SGX," in *Proceedings of the 9th Asia-Pacific Workshop on Systems*, pp. 1–8, Jeju Island, Republic of Korea, August 2018.
- [31] W. Zheng, Y. Wu, X. Wu et al., "A survey of Intel SGX and its applications," *Frontiers of Computer Science*, vol. 15, no. 3, Article ID 153808, 2021.
- [32] M. U. Sardar, R. Faqeh, and C. Fetzer, "Formal foundations for Intel SGX data center attestation primitives," in *Formal Methods and Software Engineering*, S.-W. Lin, Z. Hou, and B. Mahony, Eds., vol. 12531, pp. 268–283, Springer International Publishing, Berlin, Germany, 2020.
- [33] F. McKeen, I. Alexandrovich, A. Berenzon, C. Rozas, and H. Shafi, "Innovative instructions and software model for isolated execution," in *Proceedings of the 2nd International Workshop on Hardware and Architectural Support for Security and Privacy - HASP '13*, Tel-Aviv, Israel, June 2013.
- [34] D. Wang, F. Yi, and X. Li, "Probabilistic visual secret sharing schemes for grey-scale images and color images," *Information Sciences*, vol. 181, no. 11, pp. 2189–2208, 2011.
- [35] H.-K. Chu, C.-S. Chang, R.-R. Lee, and N. J. Mitra, "Halftone QR codes," *ACM Transactions on Graphics*, vol. 32, no. 6, pp. 1–8, 2013.
- [36] Z. Gu, W. Hu, C. Zhang, H. Lu, L. Yin, and L. Wang, "Gradient shielding: towards understanding vulnerability of deep neural networks," *IEEE Trans. Netw. Sci. Eng.*, vol. 8, no. 2, pp. 921–932, 2021.
- [37] B. Zhu, Z. Gu, Y. Qian, F. Lau, and Z. Tian, "Leveraging transferability and improved beam search in textual adversarial attacks," *Neurocomputing*, vol. 500, pp. 135–142, 2022.
- [38] M. Shafiq, Z. Tian, A. K. Bashir, A. Jolfaei, and X. Yu, "Data mining and machine learning methods for sustainable smart cities traffic classification: a survey," *Sustainable Cities and Society*, vol. 60, Article ID 102177, 2020.
- [39] M. Shafiq, Z. Tian, A. K. Bashir, X. Du, and M. Guizani, "CorrAUC: a malicious bot-IoT traffic detection method in IoT network using machine-learning techniques," *IEEE Internet of Things Journal*, vol. 8, no. 5, pp. 3242–3254, 2021.
- [40] O. Russakovsky, J. Deng, H. Su et al., "ImageNet large scale visual recognition challenge," *International Journal of Computer Vision*, vol. 115, no. 3, pp. 211–252, 2015.
- [41] M. Abadi, P. Barham, J. Chen, Z. Chen, and X. Zhang, *TensorFlow: A System for Large-Scale Machine Learning*, USENIX Association, California, CA, USA, 2016.
- [42] E. Decencière, X. Zhang, G. Cazuguel et al., "Feedback on a publicly distributed image database," *Image Analysis and Stereology*, vol. 33, no. 3, 231 pages, 2014.
- [43] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi, "Inception-v4, inception-ResNet and the impact of residual connections on learning," in *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, pp. 4278–4284, California, CA, USA, February 2017.

## Research Article

# An Explainable Statistical Method for Seizure Prediction Using Brain Functional Connectivity from EEG

Hao Chen <sup>1</sup>, Taoyun Ji <sup>2</sup>, Xiang Zhan <sup>1,3</sup>, Xiaoxin Liu <sup>3</sup>, Guojing Yu <sup>2</sup>,  
Wen Wang <sup>2</sup>, Yuwu Jiang <sup>2</sup> and Xiao-Hua Zhou <sup>1,3,4</sup>

<sup>1</sup>Beijing International Center for Mathematical Research, Peking University, No. 5 Yiheyuan Road, Haidian District, Beijing 100871, China

<sup>2</sup>Department of Pediatrics and Pediatric Epilepsy Center, Peking University First Hospital, No. 1 Xi'an Men Street, West District, Beijing 100034, China

<sup>3</sup>Department of Biostatistics, School of Public Health, Peking University, No. 38 Xueyuan Road, Haidian District, Beijing 100083, China

<sup>4</sup>Pazhou Lab, Guangzhou 510330, Guangdong, China

Correspondence should be addressed to Yuwu Jiang; [jiangyuwu@bjmu.edu.cn](mailto:jiangyuwu@bjmu.edu.cn) and Xiao-Hua Zhou; [azhou@math.pku.edu.cn](mailto:azhou@math.pku.edu.cn)

Received 18 August 2022; Revised 19 September 2022; Accepted 28 September 2022; Published 8 December 2022

Academic Editor: Inam Ullah

Copyright © 2022 Hao Chen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Background.** Epilepsy is a group of chronic neurological disorders characterized by recurrent and abrupt seizures. The accurate prediction of seizures can reduce the burdens of this disorder. Now, existing studies use brain network features to classify patients' preictal or interictal states, enabling seizure prediction. However, most predicting methods are based on deep learning techniques, which have weak interpretability and high computational complexity. To address these issues, in this study, we proposed a novel two-stage statistical method that is interpretable and easy to compute. **Methods.** We used two datasets to evaluate the performance of the proposed method, including the well-known public dataset CHB-MIT. In the first stage, we estimated the dynamic brain functional connectivity network for each epoch. Then, in the second stage, we used the derived network predictor for seizure prediction. **Results.** We illustrated the results of our method in seizure prediction in two datasets separately. For the FH-PKU dataset, our approach achieved an AUC value of 0.963, a prediction sensitivity of 93.1%, and a false discovery rate of 7.7%. For the CHB-MIT dataset, our approach achieved an AUC value of 0.940, a prediction sensitivity of 93.0%, and a false discovery rate of 11.1%, outperforming existing state-of-the-art methods. **Significance.** This study proposed an explainable statistical method, which can estimate the brain network using the scalp EEG method and use the net-work predictor to predict epileptic seizures. **Availability and Implementation.** R Source code is available at <https://github.com/HaoChen1994/Seizure-Prediction>.

## 1. Introduction

Epilepsy is a group of chronic neurological disorders characterized by the abnormal and excessive firing of brain neurons, called epileptic seizures [1]. According to the newest WHO global report on epilepsy, around 50 million people are suffering from epilepsy globally [2, 3]. During epileptic seizures, electrical activities in the brain are disrupted, resulting in dysfunction and communication disorders among brain regions, which in turn lead to many temporary symptoms, such as loss of consciousness, staring, and disturbances of movement [4]. Unpredictable seizures

dramatically affect the life of patients and may even lead to death [5]. Therefore, accurate and reliable seizure prediction can be beneficial for treating epilepsy. Patients can use Anti-Seizure Medications (ASMs) for treatment in advance, which would substantially improve the quality of life of these patients and prevent some traumatic events, including a series of life-threatening accidents.

Electroencephalography (EEG), as an electrophysiological monitoring approach to detecting brain electrical activity, has been proven to be a critical technique for diagnosing patients with epilepsy. Scalp EEG is typically noninvasive with multiple electrodes placed along the scalp [6]. It records

the spontaneous electrical activity generated by brain neurons with high time resolution over a while. Scientists have found that it can be categorized into four different waveforms for scalp EEG records of patients with epilepsy. In the view of brain functional connectivity, these four different waveforms can be represented by four brain functional connectivity structures [7], corresponding to four different states of epilepsy seizures: (1) preictal state, which is the state before a seizure occurs; (2) ictal state, which is the onset state of seizure; (3) postictal state, which refers to the immediate state after a seizure; (4) interictal state, that is the state between postictal state and preictal state [8]. Figure 1 shows the sketch of the four states. Predicting seizures can be realized by detecting the preictal state, which can be achieved through discovering the changes in brain connectivity networks from interictal state to preictal state [9–12]. However, it is clinically difficult to identify the preictal state by visual inspection of scalp EEG signals to observe changes in the structure of brain connectivity networks. Therefore, powerful and explainable statistical methods are needed to determine the preictal state from scalp EEG recordings for seizure prediction.

Nowadays, brain functional connectivity modeling approaches have been proven to be a crucial tool in the neuroscience research field [12]. The brain can be seen as a complex network in which each brain region communicates and cooperates to carry out different functions [13]. However, the dysfunction in certain areas would interfere with the processing of upcoming information, consequently leading to network disorders and changes in a person’s behavior [14]. Current research has shown that epilepsy is a specific disease related to the brain network abnormalities, and they also suggest that the brain functional connectivity of a particular patient would abnormal dynamic changes during seizures, and the forms of brain functional connectivity are different among four different states of epilepsy seizures [15]. Hence, it is reasonable and adequate to employ brain functional connectivity modeling approaches to predict seizures. Some previous studies have applied brain functional connectivity modeling methods to study epilepsy disease. For example, Williamson et al. [16] constructed multiple spatiotemporal correlation structure features from EEG data to classify the patients’ preictal or interictal states. A potential limitation of this study is that it only used a cortical network rather than a whole-brain network and cannot extract all essential features, making it difficult to achieve excellent predictive performance. Varotto et al. [17] proposed a method that employed a partially directed coherence method to depict the brain functional connectivity network. However, this study did not use the brain network features to predict seizures. Furthermore, a more recent work [18] has proposed an automatic seizure prediction method based on a graph convolutional network. This method could achieve a good seizure prediction performance by exploring the critical brain network features. To the best of our knowledge, this method is an excellent approach for seizure prediction. However, there are two fatal issues with this method. First, since this method is based on the deep learning technique, the interpretability is relatively weak. Second, the algorithm of this method is too complicated to be applied by clinicians.

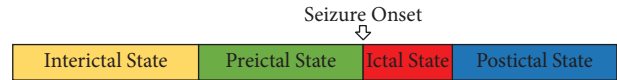


FIGURE 1: The sketch plot of four different states of epilepsy seizures.

To address the issues mentioned above, we referred to the existing novel statistical analysis framework called simultaneous differential network analysis and classification for matrix-variate data (SDNCMV) [19] and based on the characteristics of scalp EEG data and epilepsy disease, we proposed an explainable statistical model for patient-specific seizure prediction. Before introducing the proposed method, we first briefly described the SDNCMV approach. This method was a two-stage data-driven approach that deals with fMRI data. The first stage estimated each subject’s brain functional connectivity network and converted this network data into vector data for prediction in the next step. In the second stage, an ensemble prediction procedure was used to conduct the prediction results. In our study, we focused on using scalp EEG data to address patient-specific prediction problems. Since EEG data and fMRI data are in the same data format as matrix data [20, 21], we can refer to SDNCMV. However, there are still some differences between the scenarios of this study and those of Chen et al. [19] that used fMRI data to classify Alzheimer’s disease. The fMRI data is a matrix data for each subject, while the EEG data is a matrix data for each epoch [22], which is artificially generated. More specifically, since the scalp EEG data used in this study have higher temporal resolution than fMRI data and the data epochs are small, we cannot use the SDNCMV method directly. We should assume that the brain functional connectivity for each subject is time-varying and then modify the first stage of SDNCMV to estimate the dynamic brain functional connectivity to obtain better performance. More details of our method will be introduced in the following section.

The significance of this study is that we proposed an explainable statistical method, which could be used to predict seizures based on the brain functional network. In addition, the proposed method is also computationally efficient, and the results can be easily interpreted. Therefore, this method can be better applied to the clinical and is more conducive to helping more patients with epilepsy.

## 2. Materials and Methods

**2.1. Scalp EEG Data.** The scalp EEG data used in this study are obtained from Children’s Hospital Boston, the Massachusetts Institute of Technology (CHB-MIT) database [23] and Peking University First Hospital (FH-PKU) database.

The CHB-MIT database is available with open access at <https://physionet.org/physiobank/database/chbmit/>. The EEG recordings were collected from 23 children with intractable seizures, of which five males with age from 3 to 22, 17 females with age from 1.5 to 19, and one child with missing gender and age data. These recordings were grouped into 24 cases since the EEG data of patient ID chb21 was

obtained one-half year after chb01 from the same child. The sampling frequency for this database was 256 Hz, and the international 10–20 EEG electrode positions and nomenclature system was used for these recordings. The difference between two adjacent electrodes obtained the signal measurement for each electrode in this scalp EEG data. In most of the 24 cases, there are 21 unique signals, while a few cases contain less or more. Hence, only the recordings that include these 21 unique signals were selected in this study to keep the results consistent.

The FH-PKU database is a private database containing 17 patients or cases. This scalp EEG data were recorded at a sampling rate of 500 Hz and used 19 signals in the international 10–20 system. In this database, a different method was used from the CHB-MIT database to measure the signal of each electrode. This method used the signal difference between the electrode and the fixed reference electrode. Furthermore, to ensure the accuracy of the results, the physiological state of the patients in different epochs was roughly the same.

In this study, we focus on patient-specific seizure prediction performance. For each patient in these two databases, the ictal state, the period when the patient experienced seizure onset, is easily detected from raw signals by doctors. Although the preictal state is challenging to identify and there is no gold standard, based on the ictal state, the preictal state can be defined by ourselves, which is the 30-minute window before a seizure occurs and is seen as the case state in this study. However, the interictal state, which is seen as the control state, is more difficult to define; it is hard to recognize the period of the postictal state. Hence, to eliminate the noise effect of the postictal state, recordings within 2 hours after the end of the seizure are removed. The period from this time to the next preictal state is defined as the interictal state. Moreover, if the time between two seizure periods is less than 2 hours, only the first one is selected for this study. We consider all epochs as samples in this study and divide the continuous EEG data within a preictal state and interictal state into nonoverlapping 60-second epochs. To reduce computational complexity, we average the data obtained every second for each signal. Then, the data of each epoch is in matrix form with 21 or 19 columns and 60 rows for the patients from CHB-MIT or FH-PKU database.

## 2.2. Dynamic Brain Functional Connectivity Estimation.

This section introduces the procedure to estimate the individual-specific dynamic brain functional connectivity measures. For each individual, we denote  $\mathbf{X}^\gamma \in \mathbb{R}^{p \times q}$  and  $\mathbf{Y}^\phi \in \mathbb{R}^{p \times q}$  as the raw scalp EEG data matrix of  $\gamma$ -th epoch for preictal state and  $\phi$ -th epoch for the interictal state, respectively, where  $p$  represents the number of electrodes and  $q$  represents the number of time points. Based on the assumption that not every region of interest in our brain is connected, we estimate the dynamic brain functional connectivity within  $\gamma$ -th epoch for the preictal state and  $\phi$ -th epoch for the interictal state via sparse precision matrices for  $\mathbf{X}^\gamma$  and  $\mathbf{Y}^\phi$ , which estimate the strength measures of brain connectivity via partial correlations. Here we mainly focus

on the procedure of how to address the raw scalp EEG data matrix in the preictal state  $\mathbf{X}^\gamma$ , while  $\mathbf{Y}^\phi$  can be dealt with similarly.

Before introducing the detailed procedure, we follow the classical matrix normal distribution framework to define the distribution of  $\mathbf{X}^\gamma$ . Assume  $\mathbf{X}^\gamma$  follows a matrix normal distribution for each  $\gamma$ , denoted as  $\mathbf{X}^\gamma \sim \mathcal{M}\mathcal{N}(\mathbf{M}_{\mathbf{X}^\gamma}^\gamma, \Sigma_{\mathbf{X}^\gamma}^\gamma \otimes \Sigma_{\mathbf{X}_S}^\gamma)$ , where  $\Sigma_{\mathbf{X}^\gamma}^\gamma = (\Sigma_{\mathbf{X}^\gamma, ij}^\gamma) \in \mathbb{R}^{q \times q}$  and  $\Sigma_{\mathbf{X}_S}^\gamma = (\Sigma_{\mathbf{X}_S, ij}^\gamma) \in \mathbb{R}^{p \times p}$  represent the covariance matrices of  $p$  electrodes locations and  $q$  time points for  $\gamma$ -th epoch, respectively. Then, for each time point  $t$  ( $1 \leq t \leq q$ ) within the  $\gamma$ -th epoch, we have  $\mathbf{X}_t^\gamma \sim \mathcal{N}(\mathbf{M}_{\mathbf{X}_t^\gamma}^\gamma, \Sigma_{\mathbf{X}_S}^\gamma)$ . If the brain functional network is stable within each epoch, there are lots of existing approaches to estimate the sparse precision matrix  $\Omega_{\mathbf{X}_S}^\gamma = (\Sigma_{\mathbf{X}_S}^\gamma)^{-1}$  in the high-dimensional setting, such as Graphical Lasso [24] and CLIME [25]. However, in the current study of epilepsy, it is more reasonable to assume that the brain functional network is changing over time. Hence, we need to estimate the dynamic sparse precision matrix  $\Omega_{\mathbf{X}_S}^\gamma(t)$  for each time point based on time-varying covariance matrices, which can be achieved by

$$\hat{\Omega}_{\mathbf{X}_S}^\gamma(t) = \arg \min_{\Omega} \left\{ \text{Tr} \left( \hat{\Sigma}_{\mathbf{X}_S}^\gamma(t) \Omega \right) - \log |\Omega| + \lambda \|\Omega\|_1 \right\}, \quad (1)$$

where  $\hat{\Sigma}_{\mathbf{X}_S}^\gamma(t) = \sum w_{it} \mathbf{X}_i^\gamma (\mathbf{X}_i^\gamma)^T / \sum w_{it}$  is a weighted covariance matrix, and we adopt a symmetric non-negative kernel function  $K(\cdot)$  to generate the over time weights as  $w_{it} = K(|i - t|/h_n)$ . It is easy to find that this objective function is based on the Graphical Lasso and given the estimated  $\hat{\Sigma}_{\mathbf{X}_S}^\gamma(t)$ , we can use the same algorithm to solve this optimization problem. Please refer to Graphical Lasso [24] for details. In practice, to obtain a better prediction performance, we substitute  $\hat{\Sigma}_{\mathbf{X}_S}^\gamma(t)$  by  $\sum_{t=1}^q \hat{\Sigma}_{\mathbf{X}_S}^\gamma(t)/q$ , which is the average of  $\hat{\Sigma}_{\mathbf{X}_S}^\gamma(t)$  over  $q$  time points and then a unified sparse precision matrix estimation within each epoch, instead of multiple different sparse precision matrices  $\hat{\Omega}_{\mathbf{X}_S}^\gamma$ , which can be used to estimate the brain functional connectivity measures. Although, here the sparse precision matrix is same for each time points within a specific epoch, it is generated via a time varying covariance matrix, so the brain functional network can be considered dynamic. In addition, we use Gaussian kernel function and set  $h_n = n^{1/3}$  to calculate the weights  $w_{it}$  in the real application, where  $n$  is the sample size.

To sum, in this study, we adopt two symmetric matrices  $\hat{\mathbf{W}}_{\mathbf{X}_S}^\gamma = (\hat{W}_{\mathbf{X}_S, ij}^\gamma) \in \mathbb{R}^{p \times p}$  and  $\hat{\mathbf{W}}_{\mathbf{Y}_S}^\phi = (\hat{W}_{\mathbf{Y}_S, ij}^\phi) \in \mathbb{R}^{p \times p}$  to measure the dynamic brain functional connectivity strengths for  $\gamma$ -th epoch in preictal state and  $\phi$ -th epoch in interictal state, respectively. Here, we vectorize them via extracting the upper triangular elements by row for each matrix and connecting them together, and define these vectors as  $\mathbf{V}_{\mathbf{X}_S}^\gamma = \text{Vec}(\hat{\mathbf{W}}_{\mathbf{X}_S}^\gamma) \in \mathbb{R}^d$  and  $\mathbf{V}_{\mathbf{Y}_S}^\phi = \text{Vec}(\hat{\mathbf{W}}_{\mathbf{Y}_S}^\phi) \in \mathbb{R}^d$  for each epoch in different states, in which each element represent an edge in brain functional network and the dimension  $d$  is equal to  $p(p-1)/2$ . Assume there are  $n_1$  epochs in preictal state,  $n_2$  epochs in interictal state and totally  $n = n_1 + n_2$  epochs. Hence, we can define the predictor matrix used in predictive model as  $\mathbf{V} \in \mathbb{R}^{n \times p(p-1)/2}$ , where  $\mathbf{V}$  can be expressed as the stack of matrix  $\mathbf{V}_{\mathbf{X}_S}^\gamma \in \mathbb{R}^{n_1 \times p(p-1)/2}$  and matrix  $\mathbf{V}_{\mathbf{Y}_S}^\phi \in \mathbb{R}^{n_2 \times p(p-1)/2}$ . In the following of this study, we are using  $\mathbf{V}$  to serve as “Network Predictor Matrix.”

**2.3. Dynamic Brain Functional Connectivity Estimation.** Given the network predictor matrix, due to the complexity and high dimensionality of the data, we adopt a penalized and ensembled logistic regression method to predict seizures. For details, we use Lasso penalty to deal with high-dimension problem and bootstrap procedure to ensemble this penalized logistic regression. Let  $Z$  as the binary response variable and its observations are  $Z_1, \dots, Z_n$ , in which  $Z_k = 1 (k = 1, \dots, n_1)$  means corresponding observations are from preictal state and  $Z_k = 0 (k = n_1 + 1, \dots, n)$  means corresponding observations are from the interictal state. We denote  $P$  as the probability of  $Z = 1$  and  $B (b = 1, \dots, B)$  is the bootstrap times.

Now, we introduce this predictive method briefly, and for more details, please refer to Chen et al. [19]. We randomly sample  $n_1$  epochs in the preictal state and sample  $n_2$  epochs in the interictal state, respectively, with replacement. Then, we repeat the resampling  $B$  times, and for each time, we employ the high-dimensional logistic regression model. We define the  $\beta^{(b)}$  as the corresponding regression coefficients vector estimated by  $b$ -th model. If there is the coefficient in  $\beta^{(b)}$  which is not equal to 0, it indicates that the corresponding edge in the brain network is meaningful for distinguishing between the preictal state and the interictal state. Finally, after whole resampling procedure, we define the estimated coefficients vector as  $\hat{\beta}^{(b)} (b = 1, \dots, B)$  and outcome for test sample as  $\hat{P}^{(b)}$ . Hence, we use  $\hat{P}_B = \hat{P}^{(b)}/B$  to denote the proportion of a new epoch is assigned to preictal state and  $\psi_i = 1/B \sum_{b=1}^B I(\hat{\beta}_i^{(b)} \neq 0)$  to denote the weight for corresponding edge in the brain network. The greater the weight, the more important this edge is. This demonstrates the interpretability of our method.

**2.4. Performance Evaluation Measures.** The evaluation measures that we adopt for the performance of seizure prediction are Sensitivity Rate (SENS), False Discovery Rate (FDR), and Area Under Curve (AUC). The SENS measures the proportion of epochs from the preictal state with a positive result, and FDR is defined as the proportion of all epochs predicted from the preictal state, which is not. Since the values of these two criteria change with the cutoff value, we select the cutoff with the highest prediction accuracy. To avoid different cutoffs affecting performance, we also present the AUC values, which is a cutoff-independent measure and would be the most comprehensive measure.

### 3. Results

This section illustrates the results of the proposed method in seizure prediction by applying it to the CHB-MIT and FH-PKU databases.

Table 1 presents the seizure prediction results for 24 patients in the CHB-MIT database. To prove that features derived via a dynamic brain functional network contain more information and improve prediction accuracy, we convert the raw scalp EEG data matrix for each epoch to a vector and stack these  $n$  vectors into a data matrix of

dimension  $n \times pq$ . Then, we feed this data matrix into our ensemble prediction model for comparison. Furthermore, considering that the raw data may contain a small amount of information, we also combine the predictor matrix derived from brain network data (BN Data) and the predictor matrix derived from raw scalp EEG data (Raw Data) as a predictor to observe its prediction performance. The results in Table 1 show that satisfactory prediction results can be obtained by using network features, while the prediction results obtained via using raw data as input are no different from random guessing. In addition, using the combination of brain network data and raw EEG data can also achieve satisfactory prediction. However, it is still worse than using brain network data only. This is because the raw data cannot provide any valuable information for prediction, so increasing these redundant variables from raw data makes prediction performance worse.

The seizure prediction results for 17 patients from the FH-PKU database are presented in Table 2. As done in Table 1, we compare the prediction performance of our method using three kinds of input features. In Table 2, we show that network features can also achieve accurate predictions. Although the prediction result obtained using the raw data is higher than random guessing, it is still unsatisfactory. Furthermore, unlike the CHB-MIT data, in this dataset, since the raw data can provide some valuable information for prediction, it is found that the best prediction performance can be obtained using the combined data.

### 4. Discussion

In this study, we have proposed an explainable statistical method to predict epileptic seizures, which is helpful in raising the alarm before seizures. More concretely, our method uses scalp EEG data to construct a dynamic brain functional connectivity network via a time-varying precision matrix estimation approach. Then, we treat these brain functional connectivity measures as predictor variables for the ensembled prediction model. Finally, through the proposed method in this study, we can obtain accurate prediction results by using these electrodes with overactive electrical discharges as predictors. In the following, we would like to discuss the findings of this study, and at the end, we will present some future research directions.

**4.1. Relationship to Other Studies.** Our study is not the first to use scalp EEG data to extract brain connectivity signatures for seizure prediction. There have been lots of state-of-the-art approaches, such as Gemein et al. [26], Tsiouris et al. [27], and Truong et al. [28]. So far as we know, the method called STS-HGCN-AL proposed by Yang et al. [18] can achieve a better seizure prediction performance among these approaches. We speculate that if our method outperforms STS-HGCN-AL, our method will outperform all existing methods. Hence, in this study, we only compare the prediction performance of our method with the method STS-HGCN-AL in the public database CHB-MIT database. It should be noted that it is difficult to draw a direct

TABLE 1: Seizures prediction results for the patients in CHB-MIT database via our method and STS-HGCN-AL method.

Patient ID	BN data			Raw data			BN + raw data			STS-HGCN-AL		
	AUC	SENS	FDR	AUC	SENS	FDR	AUC	SENS	FDR	AUC	SENS	FDR
chb01	0.968	95.0	0.060	0.496	47.5	0.374	0.970	95.0	0.088	0.996	100	0.000
chb02	0.984	100	0.102	0.634	75.0	0.473	0.973	90.0	0.059	0.897	100	0.145
chb03	0.992	95.5	0.022	0.691	50.0	0.193	0.989	95.5	0.029	0.928	83.3	0.173
chb04	0.867	86.7	0.290	0.595	50.0	0.247	0.825	93.3	0.362	NA	NA	NA
chb05	0.835	91.5	0.173	0.458	100	0.946	0.778	60.0	0.135	0.875	100	0.000
chb06	0.804	84.3	0.344	0.505	22.9	0.154	0.779	65.7	0.205	0.906	100	0.162
chb07	0.978	93.3	0.067	0.583	43.3	0.221	0.975	93.3	0.057	NA	NA	NA
chb08	0.953	85.0	0.078	0.548	40.0	0.200	0.933	85.0	0.100	0.999	100	0.000
chb09	0.963	90.0	0.043	0.553	96.7	0.815	0.950	90.0	0.069	0.843	100	0.092
chb10	0.947	86.7	0.073	0.544	25.0	0.110	0.922	85.0	0.127	0.977	83.3	0.171
chb11	0.990	95.0	0.039	0.696	80.0	0.459	0.989	95.0	0.036	0.940	100	0.123
chb12	0.911	76.8	0.095	0.479	96.4	0.919	0.904	87.5	0.162	NA	NA	NA
chb13	0.998	96.3	0.000	0.506	22.2	0.100	0.996	96.3	0.025	0.915	85.7	0.109
chb14	0.872	95.0	0.015	0.490	42.5	0.335	0.826	92.5	0.369	0.976	100	0.104
chb15	0.836	99.8	0.231	0.433	84.6	0.838	0.793	73.6	0.306	NA	NA	NA
chb16	0.950	93.3	0.189	0.486	40.0	0.322	0.904	96.7	0.211	0.954	87.5	0.187
chb17	0.960	95.0	0.080	0.542	75.0	0.587	0.954	90.0	0.073	0.826	100	0.237
chb18	0.912	100	0.300	0.553	30.0	0.143	0.901	87.5	0.259	0.992	75.0	0.138
chb19	0.976	91.0	0.046	0.610	63.6	0.390	0.974	100	0.167	0.991	100	0.038
chb20	0.987	100	0.077	0.386	33.3	0.319	0.980	100	0.117	0.982	100	0.184
chb21	0.957	90.0	0.085	0.451	5.0	0.002	0.923	85.0	0.115	0.833	100	0.156
chb22	0.924	96.3	0.245	0.485	40.7	0.283	0.860	85.2	0.219	0.997	100	0.000
chb23	0.994	100	0.011	0.409	10.0	0.043	0.990	100	0.031	0.990	100	0.047
chb24	1.000	100	0.000	0.599	87.0	0.637	1.000	100	0.000	NA	NA	NA
Average	0.940	93.0	0.111	0.530	52.5	0.380	0.920	89.3	0.138	0.938	95.5	0.109

TABLE 2: Seizures prediction results for the patients in FH-PKU database via our method.

Patient ID	BN data			Raw data			BN + raw data		
	AUC	SENS	FDR	AUC	SENS	FDR	AUC	SENS	FDR
200002	0.950	100	0.125	0.637	40.0	0.025	0.945	80.0	0.000
210416	0.991	95.0	0.026	0.945	100	0.184	0.983	95.0	0.053
210443	0.957	90.0	0.000	0.726	70.0	0.359	0.967	90.0	0.000
210447	0.900	89.5	0.237	0.749	78.9	0.342	0.865	94.7	0.316
210454	0.891	85.0	0.105	0.834	100	0.447	0.930	90.0	0.132
210460	0.996	100	0.071	0.847	90.0	0.284	0.995	96.7	0.035
210465	0.992	95.0	0.035	0.790	80.0	0.319	0.985	100	0.113
210467	0.939	90.0	0.029	0.795	70.0	0.143	0.967	90.0	0.029
210470	0.975	90.0	0.000	0.820	100	0.447	0.978	90.0	0.000
210471	0.929	80.0	0.025	0.579	50.0	0.256	0.910	85.0	0.154
210477	0.971	87.5	0.026	0.651	95.8	0.692	0.927	87.5	0.103
210486	0.995	100	0.025	0.869	65.0	0.026	0.989	95.0	0.051
210489	0.978	90.0	0.026	0.749	85.0	0.368	0.988	90.0	0.026
210494	1.000	100	0.000	0.949	92.9	0.154	1.000	100	0.000
210498	0.992	100	0.028	0.611	70.0	0.451	0.987	100	0.056
210499	0.948	80.0	0.013	0.730	60.0	0.197	0.970	100	0.158
210503	0.877	80.0	0.040	0.981	100	0.067	0.986	100	0.080
Average	0.958	91.3	0.048	0.780	79.3	0.280	0.963	93.1	0.077

comparison due to different data preprocessing, such as how to choose the length of the epoch, when the preictal state starts, etc. In addition, the method STS-HGCN-AL has more strict requirements for the raw data, so this method can address part of patients' data within the CHB-MIT database, while our method can deal with all patients' data in terms of seizure prediction. We choose the AUC value as a measure of prediction performance to compare these two methods for the part of patients' data. From the results in Table 1, it is not

difficult to see that our method can predict more patients and has a higher AUC value of 94% in these patient data, which can suggest that our method performs better than the existing method STS-HGCN-AL in terms of predicting epileptic seizures.

Furthermore, since our method uses brain functional connectivity features to predict seizures, in this study, in addition to the prediction performance, we also briefly discuss the performance of critical feature selection. We



randomly select two patients from the FH-PKU database, the patient 210486 and patient 210494. Based on the order of the derived weights  $\psi$  for each edge in the brain network, we show the top 10 critical brain functional connectivity features identified by our model in Table 3 and Figures 2 and 3. Then, based on the weights for each edge in Table 3, we may consider the features with larger weights as identifying potential connectome biomarkers. For the identification of potential connectome biomarkers, there are many existing methods, such as Song et al. [29], Lu et al. [30], and Ding et al. [31]. However, there is no gold standard for the public datasets identifying potential connectome biomarkers. We cannot prove the effectiveness of our method for this issue and thereby cannot compare our method with existing methods. We just put the identification results here without evaluating the performance, and for this issue, we will leave it as a direction for future research.

**4.2. Limitations and Future Directions.** Although our method achieves satisfactory results in terms of seizure prediction, it still leaves much space for improvement to obtain even more realistic models. The limitations of our method proposed in this study mainly concentrated on three aspects. At first, for the sake of simplicity of our model, we have assumed that the raw scalp EEG data of each patient comes from a normal distribution. However, in the real world, we cannot know the actual distribution of the raw data. Hence, we are currently applying some statistical methods to relax the normal assumption for this problem. Secondly, in our model, we have used the same time window to define the preictal state of each patient, while in the real world, each patient has its heterogeneity and the time window of the preictal state is diverse. Hence, we plan to focus on how to estimate an optimal time window of the preictal state in the future. In the end, the study only briefly discussed how to identify potential connectome biomarkers using our method but did not test whether these biomarkers actually affected epilepsy. Therefore, in the future, we will conduct hypothesis tests in this field to demonstrate the efficiency and accuracy of our method in identifying potential connectome biomarkers.

**4.3. Potential Applications of the Method in Treatment of Epilepsy.** Epileptic seizures are sudden and have no apparent signs. The prediction of epileptic seizures can significantly enhance the effect of epilepsy treatment, improve the quality of life of patients with epilepsy, and reduce the mortality due to epileptic seizures, so the accurate prediction of epileptic seizures in the clinical application has a vital significance. Our experimental results and the comparison with previous work demonstrate that the proposed method is efficient. This gives the patient enough time to take action to cope with the seizure and reduce anxiety and trauma.

Patients with epilepsy after regular ASMs treatment, there is still one-third of patients with epilepsy that cannot be controlled. Uncontrolled seizures have severe impacts on patients' cognition, memory, quality of life, social psychology, and the growth and development of children.

TABLE 3: Epileptogenic focus localization results for the two patients in FH-PKU database.

	210486	$\psi$	210494	$\psi$
1	Occipital L $\leftrightarrow$ temporal L	1.00	Parietal L $\leftrightarrow$ occipital L	1.00
2	Frontal L $\leftrightarrow$ frontal R	0.91	Occipital L $\leftrightarrow$ frontal M	0.99
3	Frontal R $\leftrightarrow$ parietal L	0.91	Parietal R $\leftrightarrow$ parietal M	0.89
4	Parietal R $\leftrightarrow$ temporal R	0.75	Frontal M $\leftrightarrow$ frontal M	0.66
5	Parietal L $\leftrightarrow$ occipital L	0.69	Frontal L $\leftrightarrow$ parietal R	0.65
6	Temporal L $\leftrightarrow$ frontal M	0.66	Temporal L $\leftrightarrow$ parietal M	0.56
7	Parietal L $\leftrightarrow$ frontal L	0.65	Occipital R $\leftrightarrow$ parietal M	0.56
8	Parietal L $\leftrightarrow$ frontal M	0.65	Frontal R $\leftrightarrow$ frontal R	0.54
9	Parietal R $\leftrightarrow$ frontal R	0.56	Occipital L $\leftrightarrow$ occipital R	0.51
10	Frontal L $\leftrightarrow$ frontal M	0.55	Frontal L $\leftrightarrow$ frontal R	0.50

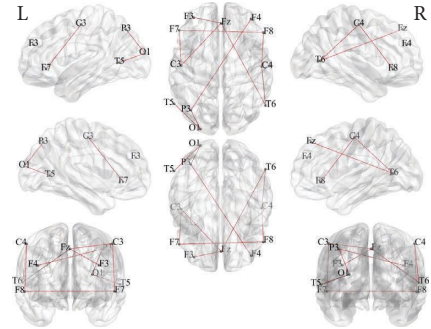


FIGURE 2: Top 10 brain functional connections that affect seizures for patient 210486.

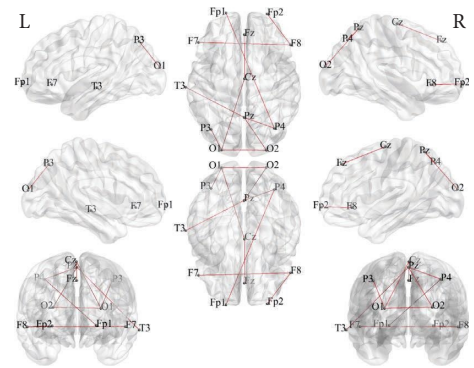


FIGURE 3: Top 10 brain functional connections that affect seizures for patient 210494.

In recent years, imaging, electroencephalography, genetics, and other diagnostic techniques have been continuously improved, and the efficacy and safety of surgical resection have been recognized. For patients with drug-resistant epilepsy with a clear epileptogenic focus and a low surgical risk, surgical resection should be considered as soon as possible. The accuracy of the connectome biomarker identification may help to determine the epileptic region before epilepsy surgery, which is the key to ensuring the success of epilepsy surgery. Hence, finding the right target remains the essential prerequisite for our new drug development, and accurate connectome biomarker identification can provide potential therapeutic targets. In this study, we have briefly discussed that our

method may enable potential connectome biomarker identification, which is essential for physicians to conduct the preoperative evaluation and develop new drugs or treatments for epilepsy, but this needs to be further validated in the future.

## 5. Consent

The parents of the patients signed written informed consent and agreed with their children's participation in this study and allowing the use of the relevant data and information for scientific research.

## Data Availability

The original contributions presented in this study are included in the article; further inquiries can be directed to the corresponding authors.

## Ethical Approval

Ethical approval for this study was obtained from the Ethics Committee of the Peking University First Hospital.

## Disclosure

Hao Chen and Taoyun Ji are the co-first authors.

## Conflicts of Interest

The authors declare that the research was conducted without any commercial or financial relationships that could be construed as a potential conflicts of interest.

## Authors' Contributions

H.C. and T.J. conceptualized the study; H.C. and X.Z. performed methodology; H.C. and X.L. accounted software; G.Y. and W.W. resourced the study; H.C. and T.J. written and prepared the original draft; X.Z. and X.L. reviewed and edited the written draft; X-H.Z. and Y.J. supervised the study; all authors have read and agreed to the published version of the manuscript.

## Acknowledgments

This research was supported by the Natural Science Foundation of China Grant (12026606); China Postdoctoral Science Foundation (2021M700249); Novo Nordisk A/S.

## References

- [1] B. S. Chang and D. H. Lowenstein, "Epilepsy," *New England Journal of Medicine*, vol. 349, no. 13, pp. 1257–1266, 2003.
- [2] H. Greter, B. Mbandando, W. Makunde et al., "Evolution of epilepsy prevalence and incidence in a tanzanian area endemic for onchocerciasis and the potential impact of community-directed treatment with ivermectin: a cross-sectional study and comparison over 28 years," *BMJ Open*, vol. 8, no. 3, Article ID e017188, 2018.
- [3] H. Yu, L. Zhu, L. Cai et al., "Variation of functional brain connectivity in epileptic seizures: an eeg analysis with cross-frequency phase synchronization," *Cognitive neurodynamics*, vol. 14, no. 1, pp. 35–49, 2020.
- [4] D. J. Englot and H. Blumenfeld, "Consciousness and epilepsy: why are complex-partial seizures complex?" *Progress in Brain Research*, vol. 177, pp. 147–170, 2009.
- [5] S. Shorvon and T. Tomson, "Sudden unexpected death in epilepsy," *The Lancet*, vol. 378, no. 9808, pp. 2028–2038, 2011.
- [6] D. Abásolo, R. Hornero, P. Espino, J. Poza, C. I. Sánchez, and R. de la Rosa, "Analysis of regularity in the eeg background activity of alzheimer's disease patients with approximate entropy," *Clinical Neurophysiology*, vol. 116, no. 8, pp. 1826–1834, 2005.
- [7] D. J. A. Smit, C. J. Stam, D. Posthuma, D. I. Boomsma, and E. J. C. De Geus, "Heritability of "small-world" networks in the brain: a graph theoretical analysis of resting-state eeg functional connectivity," *Human Brain Mapping*, vol. 29, no. 12, pp. 1368–1378, 2008.
- [8] B. Maimaiti, H. Meng, Y. Lv et al., "An overview of eeg-based machine learning methods in seizure prediction and opportunities for neurologists in this field," *Neuroscience*, vol. 481, pp. 197–218, 2022.
- [9] L. Cousyn, V. Navarro, and M. Chavez, "Preictal state detection using prodromal symptoms: a machine learning approach," *Epilepsia*, vol. 62, no. 2, pp. e42–e47, 2021.
- [10] M. Seeck, F. Cois Lazeyras, C. Michel et al., "Non-invasive epileptic focus localization using eeg-triggered functional mri and electromagnetic tomography," *Electroencephalography and Clinical Neurophysiology*, vol. 106, no. 6, pp. 508–512, 1998.
- [11] S. M. Sadjadi, E. Ebrahimzadeh, M. Shams, Ma-soud Seraji, and H. Soltanian-Zadeh, "Localization of epileptic foci based on simultaneous eeg-fmri data," *Frontiers in Neurology*, vol. 12, Article ID 645594, 2021.
- [12] P. Van Mierlo, M. Papadopoulou, E. Carrette et al., "Functional brain connectivity from eeg in epilepsy: seizure prediction and epileptogenic focus localization," *Progress in Neurobiology*, vol. 121, pp. 19–35, 2014.
- [13] O. Sporns, "The human connectome: a complex network," *Annals of the New York Academy of Sciences*, vol. 1224, no. 1, pp. 109–125, 2011.
- [14] E. Gleichgerricht, M. Kocher, and L. Bonilha, "Connectomics and graph theory analyses: novel insights into network abnormalities in epilepsy," *Epilepsia*, vol. 56, no. 11, pp. 1660–1668, 2015.
- [15] F. Li, Y. Liang, L. Zhang et al., "Transition of brain networks from an interictal to a preictal state preceding a seizure revealed by scalp eeg network analysis," *Cognitive Neurodynamics*, vol. 13, no. 2, pp. 175–181, 2019.
- [16] J. R. Williamson, D. W. Bliss, D. W. Browne, and J. T. Narayanan, "Seizure prediction using eeg spatiotemporal correlation structure," *Epilepsy and Behavior*, vol. 25, no. 2, pp. 230–238, 2012.
- [17] G. Varotto, L. Tassi, S. Franceschetti, R. Spreafico, and F. Panzica, "Epileptogenic networks of type ii focal cortical dysplasia: a stereo-eeg study," *NeuroImage*, vol. 61, no. 3, pp. 591–598, 2012.
- [18] Y. Li, Y. Yu, Y. Z. Guo, X. F. Liao, B. Hu, and T. Yu, "Spatio-temporal-spectral hierarchical graph convolutional network with semisupervised active learning for patient-specific seizure prediction," *IEEE Transactions on Cybernetics*, vol. 1, pp. 1–16, 2021.

- [19] H. Chen, Y. Guo, Y. He et al., “Simultaneous differential network analysis and classification for matrix-variate data with application to brain connectivity,” *Biostatistics*, vol. 23, no. 3, pp. 967–989, 2022.
- [20] P. Šeba, “Random matrix analysis of human eeg data,” *Physical Review Letters*, vol. 91, no. 19, Article ID 198104, 2003.
- [21] H. Hung and C. C. Wang, “Matrix variate logistic regression model with application to eeg data,” *Biostatistics*, vol. 14, no. 1, pp. 189–202, 2013.
- [22] A. Delorme, T. Sejnowski, and S. Makeig, “Enhanced detection of artifacts in eeg data using higher-order statistics and independent component analysis,” *NeuroImage*, vol. 34, no. 4, pp. 1443–1449, 2007.
- [23] H. S. Ali, *Application of machine learning to epileptic seizure onset detection and treatment*, PhD Thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, 2009.
- [24] J. Friedman, T. Hastie, and R. Tibshirani, “Sparse inverse covariance estimation with the graphical lasso,” *Biostatistics*, vol. 9, no. 3, pp. 432–441, 2008.
- [25] T. Cai, W. Liu, and Xi Luo, “A Constrained  $\ell_1$  Minimization approach to sparse precision matrix estimation,” *Journal of the American Statistical Association*, vol. 106, no. 494, pp. 594–607, 2011.
- [26] L. A. W. Gemein, R. T. Schirrmeyer, P. Chrabaszcz et al., “Machine-learning-based diagnostics of eeg pathology,” *NeuroImage*, vol. 220, Article ID 117021, 2020.
- [27] K M Tsiouris, V. C. Pezoulas, M. Zervakis, S. Konitsiotis, D. D. Koutsouris, and D. I. Fotiadis, “A long short-term memory deep learning network for the prediction of epileptic seizures using eeg signals,” *Computers in Biology and Medicine*, vol. 99, pp. 24–37, 2018.
- [28] N. D. Truong, A. D. Nguyen, L. Kuhlmann et al., “Convolutional neural networks for seizure prediction using intracranial and scalp electroencephalogram,” *Neural Networks*, vol. 105, pp. 104–111, 2018.
- [29] J. Song, D. M. Tucker, T. Gilbert et al., “Methods for examining electrophysiological coherence in epileptic networks,” *Frontiers in Neurology*, vol. 4, no. 55, p. 55, 2013.
- [30] Y. Lu, L. Yang, G. A. Worrell, and B. He, “Seizure source imaging by means of fine spatio-temporal dipole localization and directed transfer function in partial epilepsy patients,” *Clinical Neurophysiology*, vol. 123, no. 7, pp. 1275–1283, 2012.
- [31] L. Ding, G. A. Worrell, T. D. Lagerlund, and B. He, “Ictal source analysis: localization and imaging of causal interactions in humans,” *NeuroImage*, vol. 34, no. 2, pp. 575–586, 2007.

## Research Article

# The Diagnostic Value of Artificial Intelligence Ultrasound S-Detect Technology for Thyroid Nodules

Peizhen Huang,<sup>1</sup> Bin Zheng,<sup>2</sup> Mengyi Li,<sup>2</sup> Lin Xu,<sup>2</sup> Sajjad Rabbani,<sup>3</sup>  
Abdulilah Mohammad Mayet,<sup>4</sup> Chengchun Chen ,<sup>2</sup> Beishu Zhan ,<sup>1</sup> and He Jun <sup>1</sup>

<sup>1</sup>Department of Ultrasound and Imaging, Wenzhou Central Hospital, Wenzhou 325000, China

<sup>2</sup>Wenzhou Medical University, Wenzhou 325000, China

<sup>3</sup>Department of Electrical Engineering, Lahore College for Women University, LCWU, Lahore, Pakistan

<sup>4</sup>Electrical Engineering Dept, King Khalid University, Abha 61411, Saudi Arabia

Correspondence should be addressed to Chengchun Chen; [cccwzmc@126.com](mailto:cccwzmc@126.com), Beishu Zhan; [303942719@qq.com](mailto:303942719@qq.com), and He Jun; [13587978375@163.com](mailto:13587978375@163.com)

Received 6 August 2022; Revised 15 September 2022; Accepted 20 September 2022; Published 26 November 2022

Academic Editor: Ateeq Ur Rehman

Copyright © 2022 Peizhen Huang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This study aimed to evaluate the consistency of ultrasound TI-RADS classification used by sonographers with different ultrasound diagnosis experience in the diagnosis of thyroid nodules and the diagnostic value of using artificial intelligence ultrasound S-Detect technology in the differentiation of benign and malignant thyroid lesions. 100 patients who underwent ultrasound examination of thyroid masses in our hospital from June 2019 to June 2021 and were further punctured or operated on were included in the study. Pathological results were used as the gold standard to evaluate ultrasound S-Detect technology and the value of TI-RADS classification and the combined application of the two in diagnosing benign and malignant thyroid TI-RADS 4 types of nodules, and the consistency of judgments of doctors of different ages is assessed by a Kappa value. There were 128 nodules in 100 patients, 51 benign nodules, and 77 malignant nodules. For senior physicians, the sensitivity of diagnosis using TI-RADS classification combined with ultrasound S-Detect technology is 93.5%, specificity is 94.1%, and accuracy is 93.8%; for middle-aged physicians using TI-RADS classification combined with ultrasound S-Detect technology for diagnosis, the sensitivity is 89.6%, specificity is 92.2%, and accuracy is 90.6%; for junior doctors, the sensitivity of diagnosis using TI-RADS classification combined with ultrasound S-Detect technology is 83.1%, specificity is 88.2%, and accuracy is 85.1%. Regardless of seniority, the combined application of artificial intelligence ultrasound S-Detect technology and TI-RADS classification can improve the diagnostic ability of sonographers for thyroid nodules and at the same time improve the consistency of judgment among physicians, and this is especially important for radiologists.

## 1. Introduction

Athyroid nodule is one of the most common nodular lesions in adults, and its incidence is getting higher and higher. The prevalence rate of Chinese residents is as high as 18.6%, and it is more common in women, and most of them are benign nodules, only 7% of thyroid nodules tend to be malignant, but 5% of patients will change from benign to malignant without timely diagnosis and treatment [1]. The incidence of thyroid cancer increases with age, and it is the malignant tumor with the largest number of new patients among all

cancer types in recent years [2]. According to statistics from the National Cancer Research Center of the United States, there were about 64,300 new cases of thyroid cancer worldwide in 2016, and about 1,980 people died of thyroid cancer [3]. Timely and accurate detection of thyroid nodules and identification of benign and malignant thyroid nodules are of great significance to improve clinical treatment effects and prognosis of patients [4].

Fine-needle aspiration biopsy (FNAB) or surgical biopsy is the gold standard for diagnosing thyroid cancer, while for incidentally discovered thyroid nodules, ultrasound is the

most commonly used noninvasive method to differentiate benign and malignant thyroid nodules [5]. Ultrasound is a real-time imaging method, which has the advantages of no radiation and no damage, high accuracy, simple and fast operation, and low cost, but it is highly dependent on the operator and has a low repetition rate of diagnostic results [6]. With the development and application of artificial intelligence (AI) technology in the medical field, the AI ultrasound intelligent auxiliary diagnosis system based on static ultrasound images has emerged, which can greatly reduce the labor of medical workers while ensuring accuracy. However, the collection of static ultrasound images is still affected by multiple factors, while dynamic AI can perform real-time synchronous dynamic analysis of nodules from multiple levels and angles, determine the nature of nodules, and further improve diagnostic efficiency. There is no domestic report on the dynamic AI ultrasound intelligently assisted real-time diagnosis system [7–11]. The application of computer-aided diagnosis (CAD) technology “S-Detect™” can realize the qualitative and quantitative automatic analysis of ultrasound images and obtain objective, repeatable, and more accurate diagnostic results. In the latest generation of ultrasonic diagnostic equipment, S-Detect™ technology adopts the method of deep learning, which can improve the accuracy of diagnosis. At present, only a few scholars have reported the preliminary application of S-Detect™ technology in the differential diagnosis of thyroid masses [9–14]. The purpose of this study was to investigate the consistency of ultrasound TI-RADS classification in the diagnosis of thyroid nodules by sonographers with different ultrasound diagnostic experiences, to explore the diagnostic value of using artificial intelligence ultrasound S-Detect technology in the differentiation of benign and malignant thyroid lesions, as well as the clinical value of the combined diagnosis of S-Detect technology and routine thyroid ultrasonography, so as to provide sonographers with a more objective and accurate assessment tests for thyroid tumors.

## 2. Research Proposal and Object

*2.1. Research Objects.* A total of 100 patients who underwent ultrasound examination of thyroid masses in our hospital from June 2019 to June 2021 and underwent further puncture or surgery were included in the study.

### 2.1.1. Inclusion Criteria

- The patient had one or more thyroid nodules
- The patient was about to undergo FNAB or surgery
- Aged  $\geq 18$  years old
- The patient signed informed consent

*2.1.2. Exclusion Criteria.* Exclusion criteria were as follows: patients with diseases that are unfavorable to the trial or pose a threat to other participants, such as mental illness; unable to cooperate with the trial operation; pregnant or breastfeeding women; poor quality of ultrasound images that

cannot meet parameter measurement and analysis; history of thyroid surgery or history of thyroid biopsy; received radiotherapy and chemotherapy; no biopsy or surgical pathology results after examination; no clear diagnosis of benign and malignant after biopsy or surgery; simple cystic nodule, calcification; the tumor boundary that could not be identified; the underlying data were incomplete. This study was approved by the hospital’s ethics committee (K20180216), and patients signed informed consent.

*2.2. Instruments and Methods.* We used SamsungRS80A (South Korea, Seoul, Samsung Madison Co., Ltd., L3-12A linear array high-frequency probe, frequency 5–13 MHz, equipped with S-Detect intelligent detection system) high-end color Doppler ultrasound diagnostic apparatus to scan the thyroid, to understand the overall situation of the thyroid, longitudinal and transverse scanning of nodules (select 2-3 nodules with pathological results and the most suspicious of malignancy) and then analyzed the color Doppler signal of the nodule, longitudinal section and transverse section of the nodule were saved. We entered the S-Detect mode, started the S-Detect automatic analysis program to automatically draw the region of interest (ROI) on the largest longitudinal section and the largest transverse section of the thyroid nodule, and output the S-Detect diagnostic report, and the ultrasound images and evaluation report records automatically analyzed by S-Detect were stored on the hard disk for later data sorting and data analysis.

### 2.3. Analytical Method

*2.3.1. Thyroid Nodule Ultrasound TI-RADS Classification Evaluation.* All cases were routinely examined by two senior doctors with more than 10 years of experience in thyroid ultrasound diagnosis. Thyroid ultrasound TI-RADS classification assessment was performed by 6 sonographers with different experience in low, medium, and high grades and divided into three groups (there are 2 people in the low-level group, divided into low-level A and low-level B, physicians with 2 years of experience in thyroid disease ultrasound diagnosis, 2 people in the middle-level group, divided into middle-level A and middle-level senior B, physicians with 5 years of experience in thyroid disease ultrasound diagnosis, and there are 2 seniors in the senior group, divided into senior A and senior B, physicians with 10 years of experience in thyroid disease ultrasound diagnosis). According to the 2017 version of the ultrasound TI-RADS classification criteria [13], a double-blind method was used to evaluate the TI-RADS classification of thyroid nodules. The diagnostic cutoff between TI-RADS4 and TI-RADS5 was used as the diagnostic cut-off point for benign and malignant tumors, and TI-RADS4 was assessed as a possible benign tumor and TI-RADS5 was assessed as a possible malignant tumor. All sonographers were unaware of the clinical information of the cases, the number of benign and malignant cases, and the final diagnosis of the cases.

**2.3.2. Evaluation of Artificial Intelligence Ultrasound S-Detect Technology.** We activate the S-Detect automatic analysis program, make precise identification of lesions based on various characteristics of lesions, and use the ultrasonic TI-RADS classification report to issue recommended diagnostic results for thyroid nodules, which are automatically generated S-Detect diagnostic reports, judge the lesion as “probably benign” or “probably malignant” and prompt the result of “binary classification,” and finally keep a record of the images and evaluation results. If the prompting results of the two sections are different, it is regarded as “probably malignant.”

**2.3.3. Combined Diagnostic Method.** If the classification results of artificial intelligence ultrasound S-Detect technology and the sonographer’s BI-RADS classification assessment are inconsistent, when artificial intelligence ultrasound S-Detect technology diagnoses lesions as “possibly benign,” the original sonographer’s TI-RADS classification diagnostic results drops down one level; when artificial intelligence ultrasound S-Detect technology diagnoses lesions as “possibly malignant,” the original sonographer’s conventional two-dimensional ultrasound TI-RADS classification diagnostic result is upgraded by one level. When the original sonographer’s conventional two-dimensional ultrasound TI-RADS classification diagnostic result is category 3, the combined diagnostic result does not decrease. All judgment results are compared with the results of puncture or surgical pathology. With the results of puncture biopsy or surgical pathology as the gold standard, the ultrasound TI-RADS 3–5 results of physicians with different seniority levels are transformed into a benign and malignant dichotomous model, and with BI-RADS category 4 being the cutoff point for benign and malignant, possibly benign includes categories 1, 2, 3, and 4 and possibly malignant includes category 5.

**2.4. Statistical Analysis.** After sorting out all the statistical results, a database was established, and SPSS (SPSS19.0, IBM) and MedCalc19.0 statistical analysis software was used for data analysis. Linear weighted kappa coefficient analysis was used to calculate the Kappa value to evaluate the consistency of ultrasound TI-RADS classification in evaluating thyroid nodules among physicians of different seniority levels. The  $2 \times 2$  contingency table and the chi-square test were used to calculate the sensitivity, specificity, and accuracy of different senior doctors applying TI-RADS classification, artificial intelligence ultrasound S-Detect technology, and the combination of different senior doctors and S-Detect technology in differential diagnosis of benign and malignant thyroid lesions. We drew the ROC characteristic curve of different senior physicians using ultrasonic TI-RADS classification, artificial intelligence ultrasound S-Detect technology, and the combined diagnosis of thyroid lesions by different senior physicians and S-Detect technology and calculated the area under the ROC characteristic curve (AUC). The Z test was used to compare the differences

in AUC of different inspection methods. The enumeration data were expressed as a percentage, and the chi-square test of paired data was used to compare the between-group comparison between the ultrasonic features of thyroid tumors judged by artificial intelligence ultrasound S-Detect technology and the results judged by sonographers. When  $P < 0.05$ , the difference was statistically significant. Consistency assessment: 0.6–0.8 is fair, 0.8–0.9 is good, and 0.9–1.0 is excellent.

### 3. Results

**3.1. Pathological Result.** There were 128 nodules in 100 patients (among patients, 37 were female and 63 were male), of which 79 were pathologically confirmed by surgical resection, 56 were pathologically confirmed by cytology or histological biopsy, 51 were benign (19 nodular goiters, 16 thyroid adenomas, 11 Hashimoto’s thyroiditis, 5 nodular goiters with cystic degeneration), and 77 malignant nodules (43 papillary thyroid carcinoma, 22 follicular thyroid carcinoma, 10 Hashimoto’s thyroiditis complicated with papillary thyroid carcinoma, and 2 medullary thyroid carcinoma).

**3.2. Diagnostic Results of Senior Physicians and Different Diagnostic Methods.** For senior physicians, 63 benign nodules and 65 malignant nodules were diagnosed by TI-RADS classification, 57 benign nodules and 71 malignant nodules were diagnosed by ultrasound S-Detect technology, and when the two were used together, 55 benign nodules and 73 malignant nodules were diagnosed, and the kappa value was between 0.6 and 1; the sensitivity, specificity, and accuracy of TI-RADS for diagnosing thyroid nodules were 83.1%, 88.2%, and 85.1%, respectively, and  $AUC = 0.796$ , and the diagnostic sensitivity of ultrasonic S-Detect technology was 89.6%, specificity was 88.2%, accuracy was 89.1%, and  $AUC = 0.869$ . When the two were combined, diagnostic sensitivity was 93.5%, specificity was 94.1%, accuracy was 93.8%, and  $AUC = 0.923$  (Table 1, Figure 1).

**3.3. Diagnostic Results of Middle-Aged Physicians and Different Diagnostic Methods.** For middle-aged physicians, 62 benign nodules and 66 malignant nodules were diagnosed by TI-RADS classification and 54 benign nodules and 74 malignant nodules were diagnosed by ultrasound S-Detect technology. When the two were used together, 53 benign nodules and 75 malignant nodules were diagnosed, and the kappa value was between 0.6 and 1; the sensitivity, specificity, and accuracy of TI-RADS for diagnosing thyroid nodules were 80.5%, 78.4%, and 80.4%, respectively, and  $AUC = 0.713$ , and the diagnostic sensitivity of ultrasonic S-Detect technology was 83.1%, specificity was 84.3%, accuracy was 83.6%, and  $AUC = 0.826$ . When the two were combined, diagnostic sensitivity was 89.6%, specificity was 92.2%, accuracy was 90.6%, and  $AUC = 0.897$  (Table 2, Figure 2).

TABLE 1: Differences in diagnosis among senior physicians.

	TI-RADS classification		Total	Kappa	Ultrasound S-Detect		Total	Combined		Total	Kappa
	Benign	Malignant			Benign	Malignant		Benign	Malignant		
Benign	45	1	46	0.701	45	2	47	48	1	49	0.871
Malignant	18	64	82		12	69	81	7	72	79	
Total	63	65	128		57	71	128	55	73	128	
Specificity (%)	83.1				89.6			93.5			
Sensitivity (%)	88.2				88.2			94.1			
Accuracy (%)	85.1				89.1			93.8			

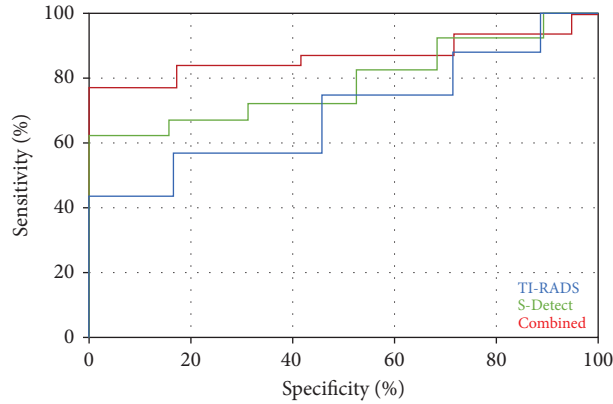


FIGURE 1: ROC curves of senior physicians using different diagnostic methods.

TABLE 2: Differences in diagnostic results among middle-aged physicians.

	TI-RADS classification		Total	Kappa	Ultrasound S-Detect		Total	Combined		Total	Kappa
	Benign	Malignant			Benign	Malignant		Benign	Malignant		
Benign	41	4	45	0.604	43	10	53	47	6	53	0.817
Malignant	21	62	83		11	64	75	8	69	75	
Total	62	66	128		54	74	128	53	75	128	
Specificity (%)	80.5				83.1			89.6			
Sensitivity (%)	78.4				84.3			92.2			
Accuracy (%)	80.4				83.6			90.6			

**3.4. Diagnostic Results of Junior Physicians and Different Diagnostic Methods.** For junior physicians, 63 benign nodules and 65 malignant nodules were diagnosed by TI-RADS classification, 61 benign nodules and 67 malignant nodules were diagnosed by ultrasound S-Detect technology, and when the two were used together, 53 benign nodules and 75 malignant nodules were diagnosed, and the kappa value was between 0.4 and 0.8; the sensitivity, specificity, and accuracy of TI-RADS for diagnosing thyroid nodules were 72.7%, 76.5%, and 74.2%, respectively, and  $AUC=0.685$ , the diagnostic sensitivity of ultrasonic S-Detect technology was 76.6%, specificity was 78.4%, accuracy was 77.3%, and  $AUC=0.798$ . When the two were combined, diagnostic sensitivity was 83.1%, specificity was 88.2%, the accuracy was 85.1%, and  $AUC=0.856$  (Table 3, Figure 3).

## 4. Discussion

The combined application of medical image data and AI based on deep learning has revolutionized the expression of medical images. Many centers have carried out static AI diagnosis of thyroid nodules in clinical practice. A large number of research data at home and abroad show that AI is effective in judging benign and malignant thyroid nodules. It has high diagnostic value, is convenient and quick, and improves clinical examination and diagnostic efficiency [8–13]. On this basis, dynamic AI uses an ultra-large-scale convolutional neural network and deep learning technology to realize real-time localization, the real-time outline of thyroid nodules, and real-time auxiliary diagnosis of benign and malignant nodules during the inspection process, making ultrasonography more efficient and precise.

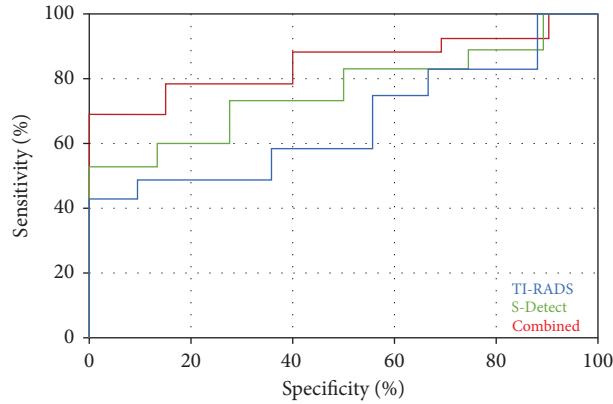


FIGURE 2: ROC curves of middle-aged physicians using different diagnostic methods.

TABLE 3: Differences in diagnostic results among junior physicians.

	TI-RADS classification		Total	Kappa	Ultrasound S-Detect		Total	Combined		Total	Kappa
	Benign	Malignant			Benign	Malignant		Benign	Malignant		
Benign	39	9	68	0.440	40	8	48	45	11	56	0.696
Malignant	14	56	70		21	59	80	8	64	72	
Total	63	65	128		61	67	128	53	75	128	
Specificity (%)	72.7				76.6			83.1			
Sensitivity (%)	76.5				78.4			88.2			
Accuracy (%)	74.2				77.3			85.1			

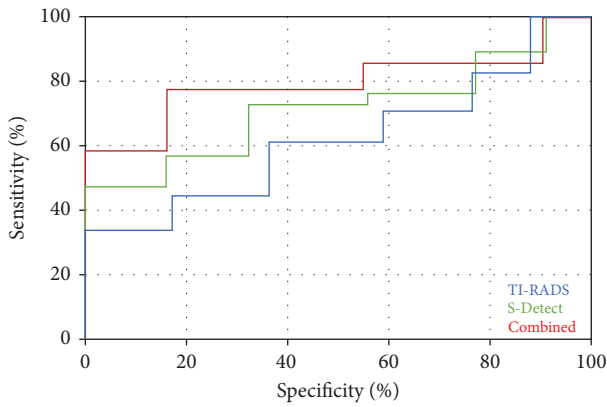


FIGURE 3: ROC curves of junior physicians using different diagnostic methods.

S-Detect<sup>TM</sup> technology is an emerging computer-aided diagnostic method, based on Korea and Russ TI-RADS (Thyroid Image Reporting and Data System) classification and ATA (American Thyroid Association) guideline classification, and using the deep learning model CNN (convolutional neural network), it automatically detects and analyzes the internal structure, echo level, boundary, direction, shape, and other information of thyroid tumors to realize the diagnosis of benign and malignant lesions [15, 16]. The application of S-Detect<sup>TM</sup> can not only improve diagnostic accuracy and repeatability but also greatly reduce the work pressure of ultrasound doctors, which has important reference value for beginners and clinicians.

Among the 128 nodules in the 100 patients in this study, 51 were benign nodules and 77 were malignant nodules confirmed by pathology. For senior physicians, the sensitivity, specificity, and accuracy of TI-RADS classification for diagnosing thyroid nodules were 89.6%, 89.6%, and 89.1%, respectively (kappa value 0.701); for middle-aged physicians, the sensitivity, specificity, and accuracy of TI-RADS classification for diagnosing thyroid nodules were 80.5%, 78.4%, and 80.4%, respectively (kappa value 0.604); for junior physicians, the sensitivity of TI-RADS classification for thyroid nodules was 72.7%, specificity was 76.5%, and accuracy was 74.2% (kappa value 0.440). With the increase in the physician seniority level, their diagnostic ability for thyroid nodules was significantly better, and they showed better evaluation consistency, confirming the influence of physician experience on the diagnosis of thyroid nodules.

Artificial intelligence ultrasound S-Detect technology is the first commercial image analysis program based on deep learning algorithms newly developed by Samsung Medical in recent years. It is installed on the high-end ultrasound diagnostic equipment of Samsung Madison in South Korea. The clinical application is mostly concentrated in the diagnosis of breast tumors [16–19], while the application in the diagnosis of thyroid nodules is less. In this study, the application of ultrasound S-Detect technology improved the specificity, sensitivity, and accuracy of thyroid nodule diagnosis among sonographers of different seniority, especially for middle and low-level doctors, and when ultrasound S-Detect technology is combined with the TI-RADS classification and the diagnostic method, it can significantly improve the diagnostic ability of middle and low-level



physicians for thyroid nodules. This is especially important for junior physicians, because their diagnostic sensitivity for benign and malignant thyroid nodules using TI-RADS classification is only 72.7%, and the specificity (76.5%) and accuracy (74.2%) also need to be improved urgently. When the ultrasound S-Detect technique was used in combination with the TI-RADS classification, the diagnostic sensitivity of junior physicians for benign and malignant thyroid nodules increased to 83.1%, which was the same as that of senior physicians using the TI-RADS classification (88.2% specificity and 85.1% accuracy), at the same time, due to the application of artificial intelligence ultrasound S-Detect technology, the judgment of physicians tends to be correct, resulting in a significant improvement in the consistency of assessment among junior physicians, from general diagnostic consistency ( $\kappa=0.44$ ) to stronger diagnostic consistency ( $\kappa=0.696$ ), and this improvement in diagnostic consistency also occurred among middle-aged and senior physicians.

In conclusion, the results of this study show that regardless of their seniority, the combined application of artificial intelligence ultrasound S-Detect technology and TI-RADS classification can improve the diagnostic ability and diagnostic consistency of sonographers for thyroid nodules. It is especially obvious for junior doctors, which helps improve the diagnostic confidence of junior sonographers and avoid unnecessary needle biopsy or surgery, which is worthy of clinical promotion. However, this study also has certain limitations; for example, it is only a single-center retrospective study, and the number of patients is small.

### Data Availability

The data used to support the findings of this study are included within the article.

### Conflicts of Interest

The authors declare that they have no conflicts of interest.

### Acknowledgments

This work was supported by the Wenzhou Basic Medical and Health Science and Technology Project (Y2020792) and the Deanship of Scientific Research at King Khalid University (Grant number RGP.1/243/42).

### References

- [1] C. D. Lansford and T. N. Teknos, "Evaluation of the thyroid nodule," *Cancer Control*, vol. 13, no. 2, pp. 89–98, 2006.
- [2] Y. Hong, X. Liu, Z. Li, X. Zhang, M. Chen, and Z. Luo, "Real-time ultrasound elastography in the differential diagnosis of benign and malignant thyroid nodules," *Journal of Ultrasound in Medicine*, vol. 28, no. 7, pp. 861–867, 2009.
- [3] J. A. Sipos, "Advances in ultrasound for the diagnosis and management of thyroid cancer," *Thyroid*, vol. 19, no. 12, pp. 1363–1372, 2009.
- [4] R. Liu, H. Li, F. Liang et al., "Diagnostic accuracy of different computer-aided diagnostic systems for malignant and benign thyroid nodules classification in ultrasound images," *Medicine*, vol. 98, no. 29, pp. 162277–e17132, 2019.
- [5] K. Kobayashi, H. Ota, M. Hirokawa et al., "Nodule in Nodule" on thyroid ultrasonography: possibility of follicular carcinoma transformed from benign thyroid tumor," *Eur Thyroid J*, vol. 6, no. 2, pp. 101–107, 2017.
- [6] M. Jang, S. M. Kim, C. Y. Lyoo, B. S. Choi, S. I. Choi, and J. H. Kim, "Differentiating benign from malignant thyroid nodules: comparison of 2- and 3- dimensional sonography," *Journal of Ultrasound in Medicine*, vol. 31, no. 2, pp. 197–204, 2012.
- [7] D. Bibicu, L. Moraru, and A. Biswas, "Thyroid nodule recognition based on feature selection and pixel classification methods," *Journal of Digital Imaging*, vol. 26, no. 1, pp. 119–128, 2013.
- [8] Q. Huang, F. Zhang, and X. Li, "Machine learning in ultrasound computer-aided diagnostic systems:A survey," *BioMed Research International*, vol. 20, no. 37, pp. 513–517, 2018.
- [9] S. Chen, W. Liu, J. Qin et al., "Research progress of computer-aided diagnosis in cancer based on deep learning and medical imaging," *Journal of Biomedical Engineering*, vol. 34, no. 2, pp. 314–319, 2017.
- [10] M. Barczyński, M. Stopa-Barczynska, B. Wojtczak, A. Czarniecka, and A. Konturek, "Clinical validation of S-DetectTM mode in semi-automated ultrasound classification of thyroid lesions in surgical office," *Gland Surgery*, vol. 9, no. S2, pp. S77–S85, 2020.
- [11] X. Wang, Y. Xie, X. Zheng et al., "A prospective multi-center randomized comparative trial evaluating outcomes of transrectal ultrasound (TRUS)-guided 12-core systematic biopsy, mpMRI-targeted 12-core biopsy, and artificial intelligence ultrasound of prostate (AIUSP) 6-core targeted biopsy for prostate cancer diagnosis," *World Journal of Urology*, vol. 14, no. 37, pp. 265–269, 2022.
- [12] W. Zheng, B. Wang, and Q. Hui, "Study on the diagnostic value of artificial intelligence technology combined with Thyroid Ultrasound Imaging and Data System (TI-RADS) grading for benign and malignant thyroid nodules," *Chinese Journal of Endocrine Surgery*, vol. 16, no. 02, pp. 185–189, 2022.
- [13] Z. Chen, "The application value of Thyroid Imaging and Reporting Data System (TI-RADS grading) combined with cervical lymph node ultrasound zoning in the diagnosis of thyroid space-occupying lesions," *Imaging Research and Medical Applications*, vol. 4, no. 9, pp. 244–245, 2020.
- [14] Q. Xia, Y. Cheng, J. Hu et al., "Differential diagnosis of breast cancer assisted by S-Detect artificial intelligence system," *Mathematical Biosciences and Engineering*, vol. 18, no. 4, pp. 3680–3689, 2021.
- [15] X. Y. Wang, L. G. Cui, J. Feng, and W. Chen, "Artificial intelligence for breast ultrasound: an adjunct tool to reduce excessive lesion biopsy," *European Journal of Radiology*, vol. 138, Article ID 109624, 2021.
- [16] H. L. Kim, E. Ju Ha, and M. Han, "Real-world performance of computer-aided diagnosis system for thyroid nodules using ultrasonography," *Ultrasound in Medicine and Biology*, vol. 45, no. 10, pp. 2672–2678, 2019.
- [17] Di Zhang, J. Fan, R. Yin et al., "A review of the role of the S-detect computer-aided diagnostic ultrasound system in the evaluation of benign and malignant breast and thyroid masses," *Medical Science Monitor*, vol. 27, no. 64, pp. 616–620, 2021.

- [18] E. Szczepanek-Parulska, K. Wolinski, K. Dobruch-Sobczak et al., "S-detect software vs. EU-TIRADS classification: a dual-center validation of diagnostic performance in differentiation of thyroid nodules," *Journal of Clinical Medicine*, vol. 9, no. 8, pp. 2495–2576, 2020.
- [19] G. Zhang and N. J. Navimipour, "A comprehensive and systematic review of the IoT-based medical management systems: applications, techniques, trends and open issues," *Sustainable Cities and Society*, vol. 82, Article ID 103914, 2022.

## Research Article

# Development and Validation of Embedded Device for Electrocardiogram Arrhythmia Empowered with Transfer Learning

Rizwana Naz Asif,<sup>1</sup> Sagheer Abbas <sup>1</sup>, Muhammad Adnan Khan <sup>2</sup>, Atta-ur-Rahman <sup>3</sup>, Kiran Sultan,<sup>4</sup> Maqsood Mahmud,<sup>5</sup> and Amir Mosavi <sup>6,7,8</sup>

<sup>1</sup>School of Computer Science, National College of Business Administration and Economics, Lahore 54000, Pakistan

<sup>2</sup>Department of Software, Gachon University, Seongnam 13120, Republic of Korea

<sup>3</sup>Department of Computer Science, College of Computer Science and Information Technology (CCSIT), Imam Abdulrahman Bin Faisal University (IAU), P.O Box 1982, Dammam 31441, Saudi Arabia

<sup>4</sup>Department of CIT, The Applied College, King Abdulaziz University, Jeddah, Saudi Arabia

<sup>5</sup>Department of Management, College of Business Administration, University of Bahrain, Zallaq, Bahrain

<sup>6</sup>Slovak University of Technology in Bratislava, Bratislava 81107, Slovakia

<sup>7</sup>Obuda University, Budapest 1034, Hungary

<sup>8</sup>TU-Dresden, Dresden 01062, Germany

Correspondence should be addressed to Muhammad Adnan Khan; [adnan@gachon.ac.kr](mailto:adnan@gachon.ac.kr) and Amir Mosavi; [amir.mosavi@kvk.uni-obuda.hu](mailto:amir.mosavi@kvk.uni-obuda.hu)

Received 2 August 2022; Revised 30 August 2022; Accepted 14 September 2022; Published 7 October 2022

Academic Editor: Ateeq Ur Rehman

Copyright © 2022 Rizwana Naz Asif et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the emergence of the Internet of Things (IoT), investigation of different diseases in healthcare improved, and cloud computing helped to centralize the data and to access patient records throughout the world. In this way, the electrocardiogram (ECG) is used to diagnose heart diseases or abnormalities. The machine learning techniques have been used previously but are feature-based and not as accurate as transfer learning; the proposed development and validation of embedded device prove ECG arrhythmia by using the transfer learning (DVEEA-TL) model. This model is the combination of hardware, software, and two datasets that are augmented and fused and further finds the accuracy results in high proportion as compared to the previous work and research. In the proposed model, a new dataset is made by the combination of the Kaggle dataset and the other, which is made by taking the real-time healthy and unhealthy datasets, and later, the AlexNet transfer learning approach is applied to get a more accurate reading in terms of ECG signals. In this proposed research, the DVEEA-TL model diagnoses the heart abnormality in respect of accuracy during the training and validation stages as 99.9% and 99.8%, respectively, which is the best and more reliable approach as compared to the previous research in this field.

## 1. Introduction

Electrocardiogram (ECG) is one of the best techniques to record the electrical signal to examine all the heart activities. If the heart is not working properly and activities are abnormal, then it would lead to serious and dangerous outcomes. Based on the World Health Organization (WHO) report, 30–40% of deaths in the entire world are due to

cardiovascular diseases which is an alarming situation, and the ratio is increasing with the passage of time. This irregular functioning and abnormalities can be found by cardiologists [1]. Literature indicates that it is difficult to identify the accurate position and transition of ECG signals with one or a simple algorithm. Artificial intelligence (AI) is playing a vital role in the recognition of heart disease by using ECG readings. The ECG value depends on the techniques,

algorithms, and different devices, which are used for detecting the signals of the ECG [2]. The rate of the heartbeat can be calculated by the QRS complex *R* peak, which represents the ECG signal per minute and shows the ventricular depolarization. The ECG wave along with its related trough and crest values such as QRS, QT, PR, and ST can explain the entire position of heart functionality [3]. The cardiologist can generate an ECG report in 3 seconds where the P wave can show atrial polarization having a duration of less than 0.12 s and amplitude of less than 2.5 mm. Then, the cardiac movement can take place from the atria to the ventricle in the PR interval of 0.12–0.2 s. The QRS complex represents the ventricular depolarization having a duration of 0.06–0.10 s. The healthy person has 0.10–0.12 s QRS, which is wider. After that, the *T* wave represents the ventricular repolarization, and it would have the same direction as the QRS. If the QRS complex is positive, the *T* wave will be positive, and if the QRS complex is negative, the *T* wave will be negative. The leading ST segment shows the interval between the depolarization and ventricular depolarization. The ST wave is a smooth wave line. So, the U wave shows the slower heart rate which is at the end leading from *T* [4]. It is important to go through the functioning of the heart and ECG before implementing it on any machine learning, deep learning, and transfer learning approach [3–5].

For the diagnosis of heart arrhythmia, the cardiologist can check the functionality of the heart and categorize it into different cardiovascular diseases upon symptoms, which can be helpful to cure the disease. The ECG analysis is one of the patterns which is applied to machine learning and deep learning to get a more precise result in a little span of time [3, 4, 6]. The study is a combination of hardware and software related to the deep learning approach. In the first part of the proposed model, build up hardware by using Raspberry Pi 3+, Arduino, a touch screen, and a heart monitor device to get the reading in real-time to create its dataset. The second part is to involve the software part, which involves different approaches of deep learning methods, and further use three databases, one from the Kaggle, i.e., Massachusetts Institute of Technology-Beth Israel Hospital (MIT-BIH Arrhythmia), the second one is own dataset named Real-Time Cardiac Arrhythmia (RT-CarArr), and the final dataset created by the combination of above two datasets, i.e., BIH-RT. The study has some parts to explain the proposed model for finding the accuracy in diagnosing ECG arrhythmia. The previous machine learning approaches to diagnose heart arrhythmia was found to be handcrafted and time-consuming. In the proposed model, the deep learning approach enhanced the accuracy and speed as compared to the machine learning methodology. In this regard, the transfer learning method AlexNet widens the research, proves the accuracy, and maximizes the results in diagnosing heart problems and abnormalities. The introduction is given in Section 1, and in Section 2, literature review, limitations, past work, and achievements are provided. Last but not the least, this will have to cover the proposed model's working performance, methodology, research tools and material, and conclusion.

## 2. Literature Review

The ECG signals which can provide the actual beat of the heart in the peak *R* in the time series have been analyzed with machine learning techniques [7]. An innovative deep learning approach and techniques help to detect abnormalities spontaneously. Deep learning made a progress in the AI field, and it could be effective for the image analysis of ECG [7, 8]. The multidimensional work (1D, 2D, and 3D) is possible with the convolutional neural network (CNN), where the 1D is limited to time series data, which are less effective as compared to 2D CNNs. So, the time series in 2D is good for machine learning algorithms [3, 9, 10]. Hence, the 2D images as the input can be applied to the ECG to make it a learning perspective and can be able to extract the features for the ECG representation. This one representation makes it possible to develop cardiovascular disease (CVD) through different automated systems. The cardiologist took some time to find the arrhythmia disease, could be an hour of observation for the analysis of ECG. The arrhythmia detection can be performed with the help of time-varying and morphological features by using hybrid feature classification. The various classification will control different kinds of wavelengths, which can occur in arrhythmia [5, 11, 12]. The IoT helped to make it possible to go through the hardware and software together with cloud computing. The ECG signal analysis worked with the classification and real-time implementation and linked with a variety of hardware (Arduino, Bluetooth, cloud servers, and phone with ECG monitor) to work parallel with the software as transfer learning and machine learning to get the accurate or required results [6, 13]. There is more work with ECG and blood pressure (BP) and their relationship; the BP value can be detected with numeric data and based on feature extraction with a machine learning model. This method was applied to estimate arterial pressure, systolic BP (SBP), and diastolic BP (DBP) by using the ECG sensor [2, 14, 15]. The classification of ECG can also be carried out with the pattern recognition method and artificial neural network (ANN) in various research studies. Moreover, the ANNs method enhanced the QRS peak detection by utilizing the multilayer perception [16, 17]. Here, the preceptor behaved like a classifier that helped to distinguish the wavelength of the normal and abnormal ECG signal reports in the form of an image to a cardiologist to read and suggest to the patient the current position either healthy or unhealthy, in the respect of heart functionality. Improvement has been seen in the ECG while using a mixture of expert (MOE) with ECG classification. In this machine learning technique, the real-time patient data can collaborate with the large dataset to get the MOE classifier. For more accuracy in the short span, the self-organizing map (SOM) is used to get the more accurate result for different kinds of heart diseases [15, 16].

Another study about the ECG has been seen in the biosignal as well; when a person has anxiety and pressure while driving a car, the supervised machine learning algorithm used the ECG movement and signals' information for biosignals and proved 72.5% accuracy [8, 16, 17]. The

detection of Alzheimer's disease uses the synchronization measures acquired with magnetoencephalography. In this way, the novel deep learning model is proposed and based on different blocks of the pooling layers, 2D convolution, and batch normalization. This model is designed to avoid overfitting, as there are massive images (25755) with few samples (132 patients). To solve the issue is to outfit the submodels with the sharing weight, and the final stage can be achieved by performing the average of submodels. Therefore, each submodel can receive the random permutation of features, which correspond to the neural activity and are arranged in the matrix form as a 2D image, which is further sorted out by a 2D convolution network. Their proposed model is a binary classifier and compared to the machine learning and deep learning approach by obtaining the best classification result with an average F1 score of 0.92 [3, 18]. Singh proposed an attention-based convolutional model to diagnose atrial fibrillation from wearable ECG. The features are extracted by using the convolution layer and classifying the atrial fibrillation. The model was tested on four databases and achieved a classification performance of 99.25% for precision, 99.25% for accuracy, and 99.50% for recall, respectively [19]. Lopez-Martin et al. presented a novel contrastive learning design and the loss function. The novel classifier was suitable for unstable and noisy datasets for intrusion detection [20].

The two main things are involved in it: one is accuracy and performance metrics and the other is computational knowledge and complexity (big O notation). The other limitation is the design and assembling techniques to operate the method or algorithm. The improvements and updating should have been seen in the datasets of the ECG. The doctors and researchers, using the portable and wearable ECG, should have to share ideas to make it real-time reliable and more effective [2, 14, 15]. The transfer learning approach can give more precise and accurate data in terms of electric ECG signals and by using MIT-DB and ECG-ID, got the satisfactory result of 97.7% for MIT-DB and 94.4% for ECG-ID and provides the qualitative result to prove the uniqueness [12, 16, 21, 22]. Jignesh et al. proposed the transfer learning of inception V3 while using the face mask for the detection of the face and achieved remarkable accuracy in testing and training [23]. The transfer learning approach is also helpful in other fields of the biomedical such as in breast cancer. Gelan Ayana et al. proposed the ImageNet transfer learning method to detect breast cancer for detection and diagnosis and achieved accuracy better as compared to the previous research [24–27]. The deep learning approach with the recurrent neural network (RNN) is useful in the ECG rhythm classifier for the sequence modeling of imbalanced data and further compared the performance of the RNN with the long short-term memory (LSTM) and gated recurrent unit (GRU) and observed that the LSTM technique is the latent method for the sequential data with an accuracy of 97.7% [28–31]. In addition, researchers proposed the validation of ECG-derived sleep architecture and ventilation in sleep apnea and chronic fatigue syndrome and analyzed the result by using the kappa score, which is 0.68, 0.85, and 0.69 for different classes [30].

Guangyu Xu proposed that the IoT built an ECG monitoring framework to improve the accuracy of the system with entire devices [22, 32–34]. An updated, comprehensive architecture for the Internet of Things devices is built on modernized blockchain models. The authors of [35] devised an intelligent way to combine IoT and blockchain in autonomous integrated sewage management. The model and framework can examine and compare various current blockchain strategies. The term “remote patient monitoring” describes keeping track of a patient's health through various digital communication channels. It uses mobile devices to collect and report on various health parameters, including those connected to the Internet of Things or the patient's body. The blockchain has been beneficial for acquiring, sharing, and storing data. It has been suggested to use IOB Health, Ethereum smart contracts, and hyperledger fabric technology [36]. Electronic health records (EHRs), electronic medical records (EMRs), remote patient monitoring, the pharmaceutical supply chain, and health insurance claims are some of the critical healthcare applications for blockchain. There is a difference between an electronic health record (EHR) and an electronic medical record, even though they are frequently used interchangeably. An EMR is similar to a digital patient chart or prescription because it records a patient's medical history and cares at a single medical facility. An electronic health record more accurately depicts a patient's overall health than it is by a paper one. Applications for EMR and EHR have been created that use and support blockchain technology Med Rec, FHIR Chain, MedShare, Ethereum applications, Med Block, and Block HIE [37]. We created an SVM-merged AlexNet classifier to handle so many attributes quickly. SVMs accelerate the hyperplane convergence in the fully connected layers of the AlexNet. Because we did not wish to begin from scratch, we used transfer learning to partially freeze the initial layers and fine-tune the features that were learned [38]. We used our architectural framework to evaluate it against the best work produced up to that point. The proposed architecture was classified more accurately than the top-ranked architectures, and its implementation took much less time. Due to this, it is a great choice when time is of the essence. The proposed algorithm [4] could be used to create better AI solutions for maternal and infant care.

In [39], the authors proposed and discussed how our design addressed typical security concerns and proposed a novel way to use blockchains to secure healthcare data. Numerous benefits exist for the proposed architecture, including increased security against known threats, decreased traffic growth, increased transparency, instantaneous traceability, and robustness. Our architecture, according to testing, reduces network traffic by a factor of 10 and speeds ledger updates by 63%. In [40], a translational combination of deep learning algorithms and CTG data was proposed, and it showed promising results with respect to accuracy and processing time. They improved the necessary time-performance metrics in medical settings. The algorithm outperformed the best architectures currently on the market with a sensitivity of 96.67%. Performance comparison of 2D and 3D CNN architectures is done for early Alzheimer's

disease symptom detection [41]. We divided people into the four groups of Alzheimer's disease (AD), non-Alzheimer's disease (NC), mild cognitive impairment (MCI), and AD using a five-fold CV method for selecting hyperparameters. Both "start from scratch" and "transfer learning" methods were used in the training of the participants. We improved the accuracy of the AD/NC classification task, the AD/MCI classification task, the NC/MCI classification task, and the AD/NC/MCI classification task using 3D CNN architectures, bringing them to 89.21%, 71.70%, 62.25%, and 59.73%, respectively. Our findings show the importance of starting from scratch in the higher domain by demonstrating that CNN architectures perform best in 3D space. The suggested forensic analysis system [42] covers IoT devices' constrained memory and resources. With the proposed forensic system, identifying the issue with Internet of Things devices in a wired environment is now simpler and faster than ever. Network traffic is sent to the logging server, where it is analyzed using previously defined rules without preventing devices from communicating. These malicious traffic logs are kept on file by the forensics server, which makes it possible for them to be recreated differently. A dataset is also produced when the Internet of Things-enabled devices record an attack. For attack detection, various machine learning models are trained and assessed. The decision tree algorithm performed admirably, with a 97.29% accuracy rate. Our plan is immediately tested when a Raspberry Pi camera is connected to the network. The decision tree's 96.01% accuracy reduced the power of machine learning models.

Some of the problems related to animal identification may be resolved with the help of the proposed research, artificial intelligence, and artificial general intelligence fields. Machine learning and federated learning are additional provider domains that could help with animal identification. The problem we just discussed might be simulated using real-time data in the future. To obtain the most accurate results, several AI-based techniques (RF, VGG-16, SVM, SMOTE, ECNN, CNN, NB, and XGBoost) were applied to various datasets. The accuracy percentages for RF, VGG, SVM, SMOTE, ECNNs, CNN, NB, and XGBoost are 98%, 97%, 92%, 90%, 88.8%, 82.15%, 81.5%, and 78.9%, respectively [43, 44]. In comparison, the machine learning and deep learning approach along with the ECG work, as well as compared to the ECG signals themselves. From the limitation perspective, machine learning needs higher computational knowledge, and the cost is high for the processors to operate [21].

*2.1. Limitations of the Related Work.* There are a few limitations regarding the previous research, as given in Table 1.

- (i) The dataset is not fused and augmented
- (ii) The new real-time dataset is not generated
- (iii) The proposed model is more accurate compared to the previous one which is comparatively showing less accuracy

- (iv) In the previous research, there is not any hardware implementation and ECG signal data have not been taken in real-time

*2.2. Our Contributions.* The major contribution of the study is as follows:

- (i) In the past research, the datasets are feature-based and handcrafted; in the proposed model, the deep learning approach (AlexNet) is applied instead of machine learning to get a more precise result.
- (ii) The real-time hardware and software are designed and implemented to get accurate results of heart arrhythmia

### 3. Proposed Model

According to the WHO, a lot of patients especially those under the age of 40 plus can have cardiovascular disease and arrest which be for a variety of reasons. The best knowledge of ML and DL with the help of different algorithms made it possible to work in more advanced ways along with the usage of innovative devices. Doctors could monitor the patients in real-time and can evaluate the sign of the diseases with the help of ECG reading with the peak values [17, 45, 46]. The research first initiates the understanding of the working of the heart and ECG electric signals and terminology to detect and diagnose cardiac arrhythmia. Three databases are used are MIT-BIT Arrhythmia with five classes (F, N, Q, S, and V), RT-CarArr with two classes (healthy and unhealthy), and BIH-RT having five classes with the combination of MIT-BIT Arrhythmia and RT-CarArr. The proposed DVEEA-TL model is comprised of two parts, namely, hardware and software. The focus is to diagnose cardiac arrhythmia by using a transfer learning model, and with the help of hardware, it is possible to get real-time images and do the IoT. Initially, the actual and main MIT-BIH Arrhythmia dataset has been taken from Kaggle, and then, the real-time dataset for the healthy and unhealthy person has been generated with the help of the proposed embedded hardware (Arduino, heart rate monitor chip, 7" touch screen with Raspberry Pi, wires, electrode pads, and so on) and able to take the reading from this 7" touch screen and later check against the proposed system if the person is healthy or unhealthy. Furthermore, the MIT-BIH dataset and the own created dataset of healthy and unhealthy patients must augment in python with the help of Keras. Then, we fused both datasets. Later, the fused dataset BIH-RT can get trained, validate, and apply performance in terms of accuracy.

*3.1. Hardware Implementation of the Proposed DVEEA-TL Model.* The emulation board and software (Arduino IDE) are compatible, in respect of hardware initiatives, and connect with the ECG sensor (AD8232) with compatible clips, cup electrodes, and ECG cables [1, 7]. Figure 1 shows the hardware connectivity for the proposed DVEEA-TL model. Furthermore, to display the ECG signal reading from the

TABLE 1: Limitations of the related work and its outcomes.

Studies	Dataset	Technique	Outcomes	Limits
Yeh et al. [1]	Private and PTB DB	ResNet, AlexNet, and SqueezeNet	Accuracy and kappa statistics of ResNet, AlexNet, and SqueezeNet in ECG waveform classification (0.97, 0.96), (0.96, 0.95), and (0.75, 0.67).	(i) No data augmentation, (ii) less accurate, and (iii) worked on waveform classification
Wasimuddin et al. [2]	ECG-ID	CAD and machine learning approach	CAD and machine learning approach working on 2D image based on classification and worked on the R peak of the ECG and showed an accuracy of 98.5%.	(i) Handcrafted features, (ii) small dataset, and (iii) accuracy is remarkable but slow because of handcrafted features
Hsu et al. [7]	MIT-BIH	AlexNet and ResNet 18	ECG into the fingerprint by using the transfer learning methods and proved the predicted accuracy of 94.4%.	(i) No data augmentation and (ii) handcrafted features
Elgendi and Menon [8]	Private	Machine learning approach	Supervised ML algorithms confirmed that ECG is an optimal wearable biosignal for assessing driving stress, with an overall accuracy of 75.02%.	(i) Low accuracy, (ii) augmentation not performed, and (iii) handcrafted features
Gaddam and Sreehari [12]	MIT-BIH	AlexNet	Transferred deep learning convolution neural net with 1D and 2D structure with 95.6% accuracy.	(i) Augmentation not performed and (ii) low accuracy
Simjanoska et al. [14]	4 private datasets used	Machine learning	The proposed method achieved 8.64 mmHg of the mean absolute error in the case of SBP.	(i) Handcrafted and (ii) low accuracy
Acharya et al. [15]	PTB DB	CNN layers	CNN for automated detection of myocardial interaction using ECG signals, and inferred the data with noise (93.5%) and without noise (95.22%).	(i) Low accuracy and (ii) less number of classes
Tomer Golany [21]	Private	GAN-based generative models such as GAN, DCNN, SIMCGAN, and SIMDCGAN.	Simulator-based network for ECG to improve deep ECG classification was used and compared all GAN-based models to find the accurate result of ECG and got SIMDCGAN as a refined and result-oriented model.	(i) Low accuracy, (ii) augmentation not performed, and (iii) handcrafted features
Sehirli et al. [28]	PTB-XL	RNN (LSTM and GRU)	Compared the performance of the RNN with the long short-term memory (LSTM) and gated recurrent unit (GRU) and then observed that the LSTM technique is the latent method for the sequential data and time series with the accuracy of 97.7%.	(i) Less accurate, (ii) small dataset, and (iii) augmentation not performed
Strothoff et al. [33]	PTB-XL	ResNet and inception	Deep learning of ECG analysis by using datasets showed an 89.8% result.	(i) No augmentation and (ii) low accuracy
Rahman et al. [45]	MIT-BIH	CAA-TL model (deep learning)	Different transfer learning approaches analyzed with data augmentation achieved 98.38% accuracy.	(i) No data fusion, (ii) low accuracy, and (iii) no hardware implementation

emulator, there is ultimate need for the monitor, and for that purpose, the Raspberry Pi 3.0+ with the 7" monitor touch screen is used which is portable and easy to operate [1, 2]. The real-time dataset RT-CarArr has been created by using the hardware, and it is comprised of 2 classes, namely, healthy and unhealthy. Whenever the whole hardware is set up, apply the three electrodes in three different places of the patient, commonly at the left arm, right arm, and left leg, as shown in Figure 1. The ECG signals will be taken from the body through electrodes and passed to AD8232, and through Arduino programming, the signals are taken and displayed on the 7" screen of Raspberry Pi 3B+. If the patient does not have

any heart problem, then it will show the frequency 360 Hz and bandwidth range from 0.5 to 40 Hz in 600 s. Finally, at this stage, the analog signals are obtained. Furthermore, the signals are compacted and reduced by fast Fourier transform (FFT) and get more refined without noise signals [1].

*3.2. Software Implementation of the Proposed DVEEA-TL Model.* The computer-based software such as python and MATLAB 2021a helped to work with the algorithm, and different deep learning approaches MATLAB and made it possible to get the prerecorded ECG data from the available

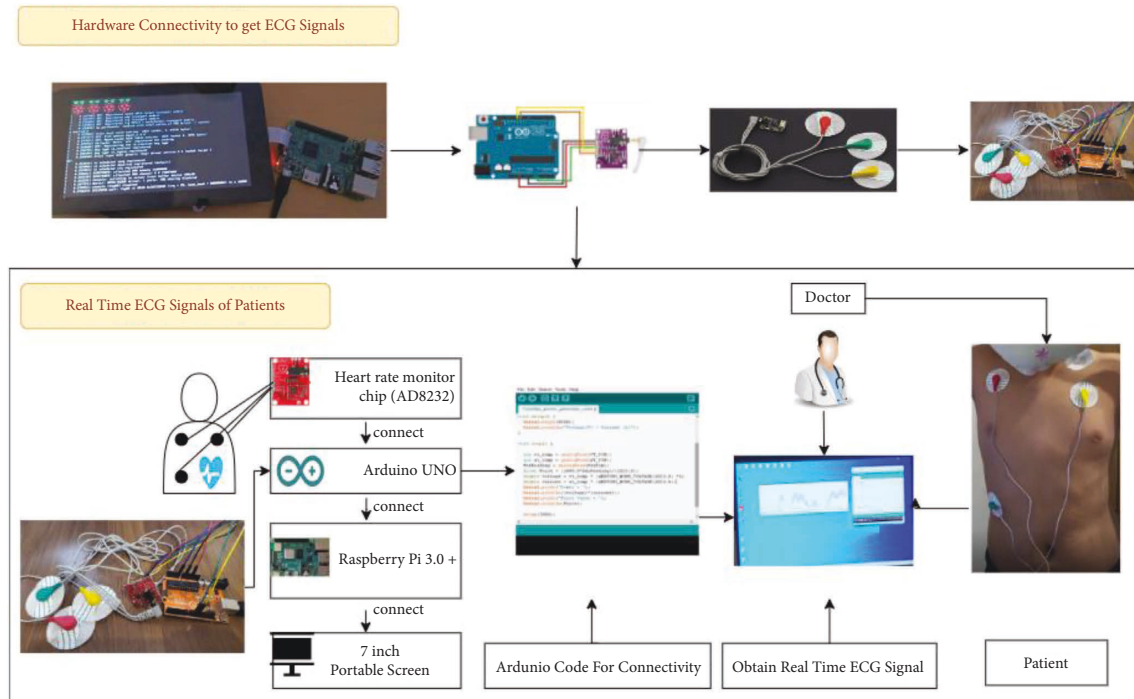


FIGURE 1: Hardware connectivity of the proposed DVEEA-TL model.

database from Kaggle and own created databases which can be preprocessed according to the required dimensions for these methods. Figure 2 shows the entire architecture of the DVEEA-TL model from hardware to software implementation with transfer learning methodology.

**3.2.1. Dataset.** As discussed earlier, all three datasets are used; one dataset is from Kaggle [47], the other is own created, and the third one is the combination of the first two datasets. The augmentation and fusion of data have been carried out against the newly generated dataset. Then, preprocess the dataset and convert it into the ECG classification by using the deep learning method AlexNet. For this purpose, the required dataset has been taken from the BIH-RT database with N (normal beat), S (supraventricular ectopic beat), V (ventricular ectopic beat), F (fusion eat), and Q (unknown beat) [47]. The number of MIT-BIH Arrhythmia for signals has been recorded for each category in this dataset as N (1500), S (3879), V (3647), F (2500), and Q (3500), respectively, which is further augmented and produce more images according to the requirement, and for the further testing, Table 2 presents the new number of samples for this MIT-BIH Arrhythmia database. In Table 3, two classes are introduced as healthy and unhealthy in the real-time database RT-CarArr with 1500 each number of images. Then, after the data fusion of both datasets, Table 4 shows the new database BIH-RT, which is generated with overall 18026 images. Furthermore, the real-time ECG images are taken with the help of hardware and extracted through frames and

compressed the signals through FFT, and then, preprocess the images. Tables 2–4 show the actual picture of all datasets which are used in the proposed DVEEA-TL model along with several classes and a number of ECG images. In the proposed DVEEA-TL model, overall, 18026 images with 5 classes are used for preprocessing, training, and validation. The whole structure of hardware and software is the best combination of innovative real-time ECG arrhythmia analysis, diagnosis, and implementation [5, 17].

**3.2.2. Transfer Learning Architecture.** The software implementation is the essential and basic requirement of ECG arrhythmia where the transfer learning method helped to find the required accuracy in diagnosing the ECG performance and find the abnormalities if found by the proposed method. Figure 3 shows the entire system in the simplest way to show the flow of working as “input data and collection,” “preprocessing,” “training,” “validation,” “performance,” and “implementation” of the proposed DVEEA-TL model.

The pseudocode of the proposed DVEEA-TL model is given in Table 5.

The deep learning approach is a widely used technique in a variety of fields such as health, transportation, agriculture, gaming, aeronautics, and so on [17]. Different pretrained transfer learning methods and models are used in this respect. Here, in the proposed DVEEA-TL model, by using AlexNet, cardiac arrhythmia can be classified and diagnosed. AlexNet is the pretrained model and has 25 layers. The images were resized according to the AlexNet parameters or



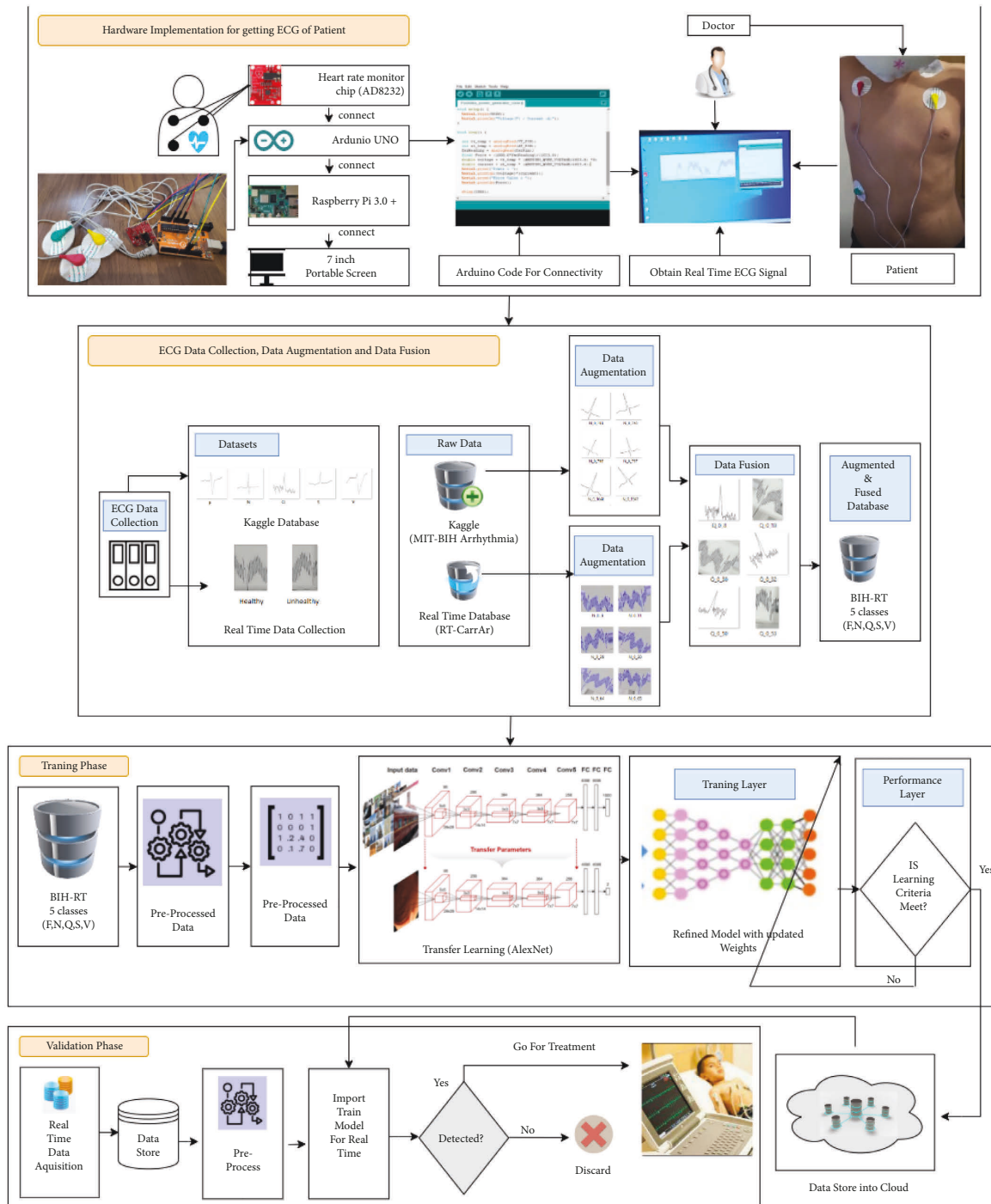


FIGURE 2: The proposed DVEEA-TL architecture.

dimensions as  $227 \times 227 \times 3$ . The preprocessing of fused data of 5 different classes shown in Figure 4 has been taken from the database BIH-RT.

After preprocessing, the fine-tuning method was applied to images, and according to the requirement, layers have been changed. Then, trained and validated all the images with a 70 : 30 ratio. The proposed DVEEA-TL model showed

99.9% with training and 99.8% with validation, respectively. Table 6 provides the number of images used for training and validation purposes.

Based on the prerequisite and the properties of the proposed DVEEA-TL model, the last three layers are changed, as shown in Figure 5. Figure 5 shows the used architecture in the proposed DVEEA-TL model. Figure 6

TABLE 2: MIT-BIH Arrhythmia augmented dataset (Kaggle) [47].

No.	Feature name	No. of samples
1	N (normal beat)	1500
2	S (supraventricular ectopic beat)	3879
3	V (ventricular ectopic beat)	3647
4	F (fusion beat)	2500
5	Q (unknown beat)	3500
	Total number of images	15026

TABLE 3: RT-CarArr augmented dataset.

No.	Feature name	No. of samples
1	Healthy (normal beat)	1500
2	Unhealthy (fusion beat)	1500
	Total number of images	3000

TABLE 4: Augmented and fused datasets BIH-RT (real-time).

No.	Feature name	No. of samples
1	N (normal beat)	3000
2	S (supraventricular ectopic beat)	3879
3	V (ventricular ectopic beat)	3647
4	F (fusion beat)	4000
5	Q (unknown beat)	3500
	Total number of images	18026

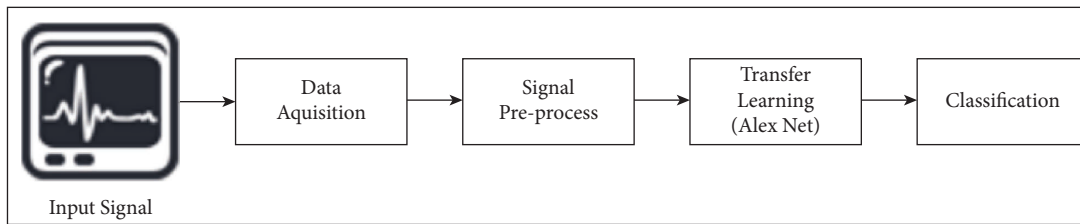


FIGURE 3: Proposed DVEEA-TL model.

TABLE 5: Pseudocode of the proposed DVEEA-TL model.

S no.	Step
1	Begin
2	Input ECG data
3	Augmentation and data fusion
4	Preprocess ECG data
5	Load data
6	Load pretrained model
7	Modified the model
8	Trained the modified model
9	Validate the modified model
10	Perform performance evaluation
11	End

shows the accuracy and loss rate of the proposed DVEEA-TL model.

#### 4. Simulation and Results

A matrix comprises accuracy, classification miss rate, sensitivity, precision, false positive ratio, false negative ratio, F1 score, Mathew correlation coefficient (MCC) analysis, specificity, and kappa score are used to evaluate the overall performance of the fine-tuning approach. The assessment

and development of the entire program are to be carried out in the MATLAB 2021a with 11th Gen Intel(R) Core (TM) i5-1135G7 @ 2.40 GHz computer processor, 8.00 GB RAM, and 1 TB hard disk along with also tested on Raspberry Pi 3.0+, Arduino, and heart monitor on run time. Performance evaluation of algorithms is evaluated with different statistical parameters as shown in the following equations adapted from [27–33, 45]:

$$\text{Accuracy} = \frac{O_{ri} + O_{rk}/I_{ri} + I_{ik}}{O_{ri}/I_{ri} + \sum_{j=1}^n (O_{ri}, j \neq i)/I_{rj} + O_{rk}/I_{rk} + \sum_{l=1}^n (O_{rl}, l \neq k)/I_{rk}}, \text{ where } l, j, i, k = 1, 2, \dots, m, \quad (1)$$

$$\text{Miss rate} = \frac{\sum_{l=1}^n O_{rl, l \neq k}/I_{rk}}{\sum_{l=1}^n (O_{rl, l \neq k})/I_{rk} + O_{ri}/I_{ri}}, \text{ where } l, j, i, k = 1, 2, \dots, m, \quad (2)$$

$$\text{True positive rate/recall} = \frac{O_{ri}/I_{ri}}{O_{ri}/I_{ri} + \sum_{l=1}^n (O_{rl}, l \neq k)/I_{rk}}, \text{ where } l, j, i, k = 1, 2, \dots, m, \quad (3)$$

$$\frac{\text{True negative rate}}{\text{Sensitivity}} = \frac{O_{rk}/I_{rk}}{O_{rk}/I_{rk} + \sum_{j=1}^n (O_{rj}, j \neq 1)/I_{rj}}, \text{ where } k, j = 1, 2, \dots, m, \quad (4)$$

$$\text{Precision} = \frac{O_{ri}/I_{ri}}{O_{ri}/I_{ri} + \sum_{j=1}^n (O_{rj}, j \neq 1)/I_{rj}}, \text{ where } j, i = 1, 2, \dots, m, \quad (5)$$

$$\text{False omission rate} = \frac{\sum_{l=1}^n O_{rl, l \neq k}/I_{rk}}{\sum_{l=1}^n (O_{rl, l \neq k})/I_{rk} + O_{rk}/I_{rk}}, \text{ where } l, k = 1, 2, \dots, m, \quad (6)$$

$$\text{False discovery rate} = \frac{\sum_{j=1}^n O_{rj, j \neq i}/I_{rj}}{O_{ri}/I_{ri} + \sum_{j=1}^n (O_{rj}, j \neq i)/I_{rj}}, \text{ where } i, j = 1, 2, \dots, m, \quad (7)$$

$$\text{F 0.5 score} = 1.25 \times \text{precision} \times \frac{\text{recall}}{0.25 \times \text{precision} + \text{recall}}, \quad (8)$$

$$\text{F1 score} = 2 \times \text{precision} \times \frac{\text{recall}}{\text{precision} + \text{recall}}, \quad (9)$$

$$\text{MCC} = \frac{(\text{TP} \times \text{TN}) - (\text{FP} \times \text{FN})}{\sqrt{(\text{TP} + \text{FP})(\text{TP} + \text{FN})(\text{TN} + \text{FP})(\text{TN} + \text{FN})}}, \quad (10)$$

$$\text{Kappa score} = \frac{p_o - p_e}{1 - p_e} = 1 - \frac{1 - p_o}{1 - p_e}. \quad (11)$$

MCC is used to measure the microaveraging for each class and then calculate the statistics. The MCC values used the true positive (TP), true negative (TN), false positive (FP), and false negative (FN) for each class, and the final statistic value is given in Table 7. With the help of the kappa score, the classification performance can be measured by using the metrics, and it can show the observed and

estimated value and the probability of its existence. Kappa score is always equal to “1” or less than “1.” The values against kappa analysis are given in Table 8. The performance of five different classes is observed by using the accuracy, sensitivity, classification miss rate, and specificity. The confusion matrix in Figures 7 and 8 shows the class-level comparison applied to transfer learning which is



FIGURE 4: Samples of 5 classes after preprocessing.

TABLE 6: The proposed DVEEA-TL model used data division during training and validation.

Classes	Total no. of instances (100%)	Training instances (70%)	Validation instances (30%)
Q	3647	2552	1095
N	3879	2715	1164
F	3000	2100	900
V	3500	2450	1050
S	4000	2800	1200
Total	18026	12617	5409

```

layers =
25x1 Layer array with layers:
 1 '' Image Input 227x227x3 images with 'zerocenter' normalization
 2 'conv1' Convolution 96 11x11x3 convolutions with stride [4 4] and padding [0 0 0 0]
 3 'relu1' ReLU ReLU
 4 'norm1' Cross Channel Normalization cross channel normalization with 5 channels per element
 5 'pool1' Max Pooling 3x3 max pooling with stride [2 2] and padding [0 0 0 0]
 6 'conv2' Grouped Convolution 2 groups of 128 5x5x48 convolutions with stride [1 1] and padding [2 2 2 2]
 7 'relu2' ReLU ReLU
 8 'norm2' Cross Channel Normalization cross channel normalization with 5 channels per element
 9 'pool2' Max Pooling 3x3 max pooling with stride [2 2] and padding [0 0 0 0]
10 'conv3' Convolution 384 3x3x256 convolutions with stride [1 1] and padding [1 1 1 1]
11 'relu3' ReLU ReLU
12 'conv4' Grouped Convolution 2 groups of 192 3x3x192 convolutions with stride [1 1] and padding [1 1 1 1]
13 'relu4' ReLU ReLU
14 'conv5' Grouped Convolution 2 groups of 128 3x3x192 convolutions with stride [1 1] and padding [1 1 1 1]
15 'relu5' ReLU ReLU
16 'pool5' Max Pooling 3x3 max pooling with stride [2 2] and padding [0 0 0 0]
17 'fc6' Fully Connected 4096 fully connected layer
18 'relu6' ReLU ReLU
19 'drop6' Dropout 50% dropout
20 'fc7' Fully Connected 4096 fully connected layer
21 'relu7' ReLU ReLU
22 'drop7' Dropout 50% dropout
23 '' Fully Connected 5 fully connected layer
24 '' Softmax softmax
25 '' Classification Output crossentropyex

```

FIGURE 5: Transfer learning architecture of the proposed DVEEA-TL.

used in architecture. All the classes showed different values as the training set values for F are 899 TP, 4 FN, 1 FP, and 4554 TN with an overall accuracy is 99.91%, as shown in Figure 7. The matrix value for N is 1164 TP, 3 FN, 0 FP, and 4290 TN with an overall accuracy is 99.94%.

Hereafter, the confusion matrix value for Q is 1094 TP, 0 FN, 6 FP, and 4358 TN with an overall accuracy is 99.89%. So, the matrix value for S is 1200 TP, 0 FN, 0 FP, and 4258 TN with an overall accuracy of 100%. Last, the matrix value for Q is 1094 TP, 0 FN, 0 false positive, and 4364 true

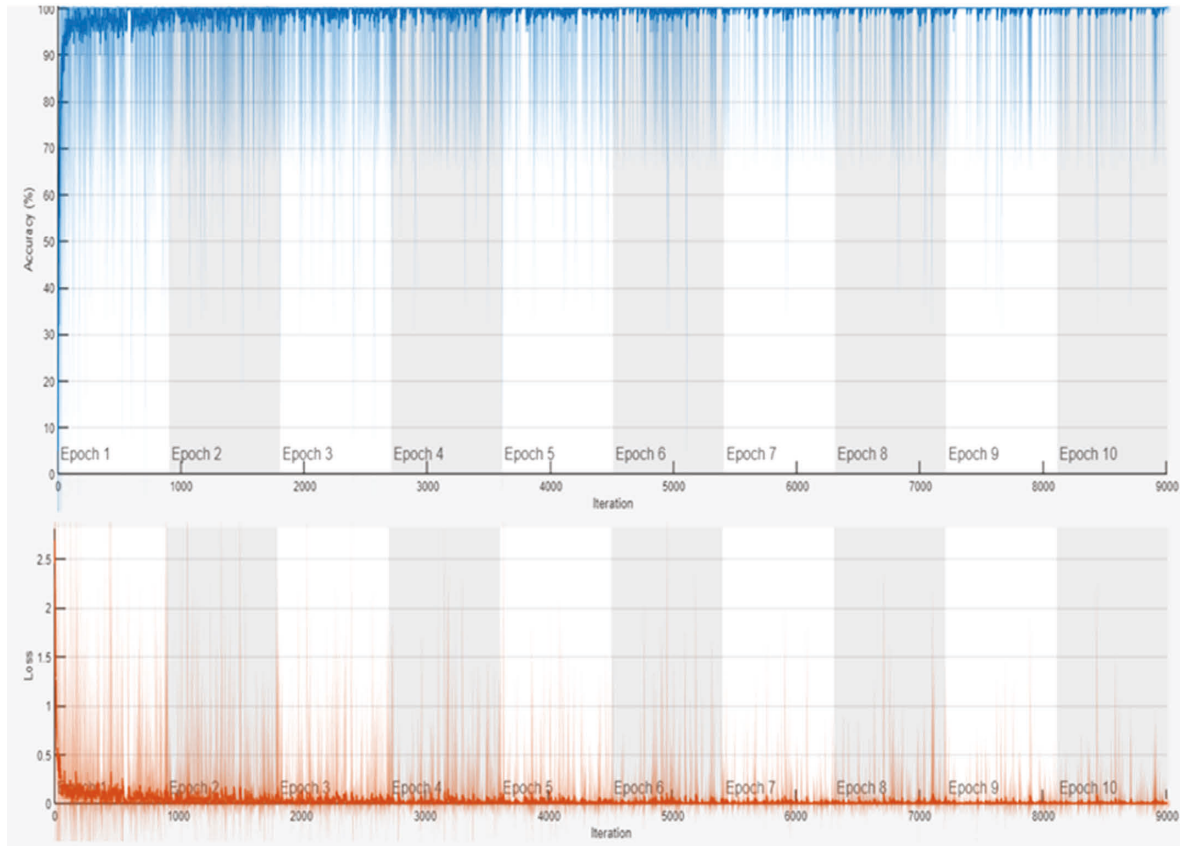


FIGURE 6: Accuracy and loss rate of the proposed DVEEA-TL model.

TABLE 7: Class-wise training and validation results of the proposed DVEEA-TL model.

T	Evaluation matrix	F	N	V	Q	S
		Fusion beat (%)	Normal beat (%)	Ventricular beat (%)	Unknown beat (%)	Supraventricular beat (%)
Accuracy	Training	99.91	99.95	99.88	100	100
	Validation	99.81	99.78	99.82	100	100
Classification miss rate	Training	0.09	0.05	0.12	0	0
	Validation	0.19	0.22	0.18	0	0
Sensitivity	Training	99.56	99.74	100	100	100
	Validation	98.90	100	100	100	100
Specificity	Training	19.74	27.13	25.07	25.07	28.19
	Validation	20.03	34.93	23.69	25.37	28.55
Precision	Training	99.89	100	54.70	100	100
	Validation	100	100	99.04	100	100
FPR	Training	0.80	0.73	0.75	0.75	0.72
	Validation	0.80	0.65	0.76	0.75	0.71
FNR	Training	0.004	0.002	0	0	0.004
	Validation	0.01	0	0	0	0

negative as shown in Figure 8. The simulation result of all classes of transfer learning is given in Tables 7 and 8, and it represents the accuracy, classification miss rate, sensitivity, specificity, precision, false negative ratio (FNR), and false positive ratio (FPR) of all five ECG classes.

### 5. Comparative Analysis of the Proposed Model

Table 9 provides the comparison result of the proposed DVEEA-TL model with the literature. This proposed model is a combination of hardware and software, which made it

TABLE 8: Proposed DVEEA-TL model' overall results.

Performance matrices	Training (%)	Validation (%)
Accuracy	99.9	99.8
Classification miss rate	0.05	0.07
Sensitivity	99.8	99.7
Specificity	21.09	26.5
Precision	90.9	99.80
F1 score	0.98	0.97
FPR	0.75	0.73
FNR	0.002	0.002
MCC	99.2	98.5
Kappa score	0.98	0.97

Confusion Matrix

Output Class	F	899 16.6%	0 0.0%	0 0.0%	0 0.0%	4 0.1%	99.6% 0.4%
	N	1 0.0%	1164 21.5%	0 0.0%	0 0.0%	2 0.0%	99.7% 0.3%
	Q	0 0.0%	0 0.0%	1094 20.2%	0 0.0%	0 0.0%	100% 0.0%
	S	0 0.0%	0 0.0%	0 0.0%	1200 22.2%	0 0.0%	100% 0.0%
	V	0 0.0%	0 0.0%	0 0.0%	0 0.0%	1044 19.3%	100% 0.0%
		99.9% 0.1%	100% 0.0%	100% 0.0%	100% 0.0%	99.4% 0.6%	99.9% 0.1%
	F	N	Q	S	V		

Target Class

FIGURE 7: Confusion matrix (training of ECG dataset).

Confusion Matrix

Output Class	F	270 16.7%	0 0.0%	0 0.0%	0 0.0%	3 0.2%	98.9% 1.1%
	N	0 0.0%	349 21.5%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	Q	0 0.0%	0 0.0%	328 20.2%	0 0.0%	0 0.0%	100% 0.0%
	S	0 0.0%	0 0.0%	0 0.0%	360 22.2%	0 0.0%	100% 0.0%
	V	0 0.0%	0 0.0%	0 0.0%	0 0.0%	311 19.2%	100% 0.0%
		100% 0.0%	100% 0.0%	100% 0.0%	100% 0.0%	99.0% 1.0%	99.8% 0.2%
	F	N	Q	S	V		

Target Class

FIGURE 8: Confusion matrix (validation of ECG dataset).

TABLE 9: Proposed DVEEA-TL model compared with the state-of-the-art literature.

Studies	Hardware implementation	Data augmentation	Data fusion	Datasets	Method	Findings
Yeh et al. [1]	Yes	No	No	PTB DB	ResNet, AlexNet, and SqueezeNet	Predicted accuracies: 97%, 95%, and 75%
Wasimuddin et al. [2]	No	No	No	ECG-ID	CAD and machine learning	Predicted accuracy: 98.5%
Vijayakumar et al. [6]	No	No	No	No	Feature extraction to remove noise	Predicted accuracy: 96.5%
Hsu et al. [7]	No	No	No	MIT-DB	AlexNet and ResNet	Predicted accuracy: 94.4%
Gaddam and Sreehari [12]	No	No	No	MIT-DB	AlexNet	Predicted accuracy: 95.6%
Simjanoska et al. [14]	No	No	No	PTB DB	ML-train-validation-test evaluation	Predicted accuracy: 98%
Acharya et al. [15]	No	No	No	PTB DB	CNN layers	Predicted accuracy: 93.5% (for noise data) 95.22% (for non-noise data)
Hammad et al. [17]	No	No	No	PTB	ResNet model	Predicted accuracy: 98.85%
Golany et al. [21]	No	No	No	MIT-DB	GAN-based model	Predicted accuracy: 97.5%
Sehirli et al. [28]	No	No	No	PTB-XL	RNN (LSTM and GRU)	Predicted accuracy: 97.7%
Strothoff et al. [33]	No	No	No	PTB-XL	ResNet and inception	Predicted accuracy: 89.8%
Rahman et al. [45]	No	Yes	No	MIT-BIH	CAA-TL model (deep learning)	Predicted accuracy: 98.38%
Proposed DVEEA-TL model	Yes	Yes	Yes	BIH-RT (real-time dataset)	Transfer learning (AlexNet)	Training (99.9%) Validation (99.8%)

distinct from the previous research. Furthermore, the real-time own dataset, data fusion, and data augmentation are the achievements of this study as well. In the comparative analysis, it observed that the proposed DVEEA-TL model is giving the reckless and most trustworthy result as compared to previously published approaches.

## 6. Conclusions

In the proposed model, ECG devices, the algorithm, dataset, and ecological and financial factors all play an important role in determining the efficiency of ECG analysis. The more critical thing in heart arrhythmia is to diagnose in the early stages to save the life, and the ECG is the best step or ointment to check the functioning of the heart ECG signals. In the proposed DVEEA model, five different classes have been classified, preprocessed, trained, and validated in the knowledge of the artificial network. Furthermore, the augmented and fusion of data improved the probability of accuracy. The proposed model DVEEA-TL has the combination of hardware along with the software in MATLAB 2021a, Python, and Arduino working, and the datasets are trained on 10 epochs. Working on different layers helped to diagnose the ECG arrhythmia that gives a 99.9% and 99.8% training accuracy and validation accuracy, respectively, which is an excellent and outstanding result for the life-threatening cardiac disease. The proposed DVEEA-TL model showed remarkable accuracy, but there are still some variations and limitations, which must be in consideration

and work in the future. For this model, the computation processing is high and consumes time to train the datasets of the ECG images. Furthermore, in the future, we can improve the computation processing to use the AWS or GPU, instead of the CPU to get the training at an enormous speed. An innovative and secure federated deep learning approach can be applied further to the proposed model to make it more consistent and steady in the medical sciences.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Disclosure

Rizwana Naz Asif and Sagheer Abbas are the co-first authors.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

The authors would like to express their deepest gratitude to Dr. Dhiaa Musleh and Dr. Suleiman Ali Alsaif from IAU for the helpful suggestions during data collection and result interpretation.

## References

- [1] L. R. Yeh, W. C. Chen, H. Y. Chan et al., "Integrating ECG monitoring and classification via IoT and deep neural networks," *Biosensors*, vol. 11, no. 6, p. 188, 2021.
- [2] M. Wasimuddin, K. Elleithy, A. S. Abuzneid, M. Faezipour, and O. Abuzagheh, "Stages-based ECG signal analysis from traditional signal processing to machine learning approaches: a survey," *IEEE Access*, vol. 8, pp. 177782–177803, 2020.
- [3] "Physical Therapy Reviewer How to Read an ECG," 2015, <https://ptreviewer.com/electrocardiogram-ecg-2/reading-an-ecg/>.
- [4] "MSD Manual Consumer Version Figure: ECG: Reading the Waves," 2021, [https://www.msmanuals.com/en-jp/home/multimedia/figure/cvs\\_ecg\\_reading](https://www.msmanuals.com/en-jp/home/multimedia/figure/cvs_ecg_reading).
- [5] I. Hussain and S. J. Park, "Big-ECG: cardiographic predictive cyber-physical system for stroke management," *IEEE Access*, vol. 9, pp. 123146–123164, 2021.
- [6] T. Vijayakumar, R. Vinothkanna, and M. Duraipandian, "Fusion based feature extraction analysis of ECG signal interpretation—a systematic approach," vol. 3, pp. 1–16, 2021.
- [7] P. Y. Hsu, P. H. Hsu, and H. L. Liu, "Fold electrocardiogram into a fingerprint," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pp. 828–829, Seattle, WA, USA, July, 2020.
- [8] M. Elgendi and C. Menon, "Machine learning ranks ECG as an optimal wearable bio-signal for assessing driving stress," *IEEE Access*, vol. 8, pp. 34362–34374, 2020.
- [9] G. T. Reddy and N. Khare, "Heart disease classification system using optimised fuzzy rule based algorithm," *International Journal of Biomedical Engineering and Technology*, vol. 27, no. 3, pp. 183–202, 2018.
- [10] S. Poudel, Y. J. Kim, D. M. Vo, and S. W. Lee, "Colorectal disease classification using efficiently scaled dilation in convolutional neural network," *IEEE Access*, vol. 8, pp. 99227–99238, 2020.
- [11] M. K. Moridani and M. Pouladian, "A novel method to ischemic heart disease detection based on non-invasive ECG imaging," *Journal of Mechanics in Medicine and Biology*, vol. 19, no. 03, pp. 1950002–1950009, 2019.
- [12] P. G. Gaddam and R. V. Sreehari, "Automatic classification of cardiac arrhythmias based on ECG signals using transferred deep learning convolution neural network in," *Journal of Physics: Conference Series*, vol. 2089, Article ID 012058, 2021.
- [13] M. K. Hasan, M. Shafiq, S. Islam et al., "Lightweight Cryptographic Algorithms for Guessing Attack protection in Complex Internet of Things Applications," *Complexity*, vol. 2021, no. 1, 13 pages, 2021.
- [14] M. Simjanoska, M. Gjoreski, M. Gams, and A. Madevska Bogdanova, "Non-Invasive Blood pressure estimation from ECG using machine learning techniques," *Sensors*, vol. 18, no. 4, pp. 1160–1172, 2018.
- [15] U. R. Acharya, H. Fujita, S. L. Oh, Y. Hagiwara, J. H. Tan, and M. Adam, "Application of deep convolutional neural network for automated detection of myocardial infarction using ECG signals," *Information Sciences*, vol. 415–416, pp. 190–198, 2017.
- [16] K. G. Celebiler, "Heart Disease Diagnosis Using Neural Networks on Electrocardiogram Datasets," 2021.
- [17] M. Hammad, P. Plawiak, K. Wang, and U. R. Acharya, "ResNet-Attention model for human authentication using ECG signals," *Expert Systems*, vol. 38, no. 6, pp. e12547–e12556, 2021.
- [18] M. Lopez-Martin, A. Nevado, and B. Carro, "Detection of early stages of Alzheimer's disease based on MEG activity with a randomized convolutional neural network," *Artificial Intelligence in Medicine*, vol. 107, pp. 101924–101931, 2020.
- [19] P. Singh, "Attention Based Convolutional Denoising Autoencoder for Two-lead ECG Denoising and Arrhythmia Classification," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, 2022.
- [20] M. Lopez-Martin, A. Sanchez-Esguevillas, J. I. Arribas, and B. Carro, "Supervised contrastive learning over prototype-label embeddings for network intrusion detection," *Information Fusion*, vol. 79, pp. 200–228, 2022.
- [21] T. Golany, K. Radinsky, and D. S. G. Freedman, *Simulator-based Generative Adversarial Networks for ECG Synthesis to Improve Deep ECG Classification in International Conference on Machine Learning*, pp. 3597–3606, 2020.
- [22] G. A. Tadesse, T. Zhu, Y. Liu et al., "Cardiovascular Disease Diagnosis Using Cross-Domain Transfer Learning," in *Proceedings of the 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 4262–4265, IEEE, Berlin, Germany, July, 2019.
- [23] G. Jignesh Chowdary, N. S. Punna, S. K. Sonbhadra, and S. Agarwal, *Face Mask Detection Using Transfer Learning of Inceptionv3 in International Conference on Big Data Analytics*, pp. 81–90, Springer, Cham, 2020.
- [24] G. Ayana, K. Dese, and S. W. Choe, "Transfer learning in breast cancer diagnoses via ultrasound imaging," *Cancers*, vol. 13, no. 4, pp. 738–746, 2021.
- [25] M. M. K. Al-Mozani, "Detection of QRS Complex and Classification of Electrocardiogram Signals Using Computational Intelligent Algorithms (Doctoral Dissertation)," *International Journal of Electrical and Computer Engineering*, vol. 9, no. 5, pp. 3512–3521, 2021.
- [26] S. Safdar, M. Rizwan, T. R. Gadekallu et al., "Bio-Imaging-based machine learning algorithm for breast cancer detection," *Diagnostics*, vol. 12, no. 5, pp. 1134–1146, 2022.
- [27] S. Mehmood, T. M. Ghazal, M. A. Khan et al., "Malignancy detection in lung and colon histopathology images using transfer learning with class selective image processing," *IEEE Access*, vol. 10, pp. 25657–25668, 2022.
- [28] E. Sehirli and M. K. Turan, "A novel method for segmentation of QRS complex on ECG signals and classify cardiovascular diseases via a hybrid model based on machine learning," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 9, no. 1, pp. 12–21, 2021.
- [29] S. R. Tithi, A. Aktar, and F. Aleem, *Machine Learning Approach for ECG Analysis and Predicting Different Heart Diseases*, Doctoral Dissertation, BRAC University, Dhaka, 2018.
- [30] J. Werth, M. Radha, P. Andriessen, R. M. Aarts, and X. Long, "Deep learning approach for ECG-based automatic sleep state classification in preterm infants," *Biomedical Signal Processing and Control*, vol. 56, pp. 101663–101678, 2020.
- [31] M. J. Decker, S. Eyal, Z. Shinar et al., "Validation of ECG-derived sleep architecture and ventilation in sleep apnea and chronic fatigue syndrome," *Sleep and Breathing*, vol. 14, no. 3, pp. 233–239, 2010.
- [32] G. Xu, "IoT-assisted ECG monitoring framework with secure data transmission for health care applications," *IEEE Access*, vol. 8, pp. 74586–74594, 2020.
- [33] N. Strodthoff, P. Wagner, T. Schaeffter, and W. Samek, "Deep learning for ECG analysis: benchmarks and insights from PTB-XL," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 5, pp. 1519–1528, 2021.
- [34] S. Y. Siddiqui, A. Athar, M. A. Khan et al., "Modelling, simulation and optimization of diagnosis cardiovascular



- disease using computational intelligence approaches,” *Journal of Medical Imaging and Health Informatics*, vol. 10, no. 5, pp. 1005–1022, 2020.
- [35] G. B. Janardhana Swamy, D. R. Janardhana, C. P. Vijay, and V. Ravi, *Blockchain-Enabled IoT Integrated Autonomous Sewage Management System in Convergence of Internet of Things and Blockchain Technologies*, pp. 41–56, Springer, Cham, Switzerland, 2022.
- [36] U. Islam, A. Muhammad, R. Mansoor et al., “Detection of distributed denial of service (DDoS) attacks in IOT based monitoring system of banking sector using machine learning models,” *Sustainability*, vol. 14, p. 8374, 2022.
- [37] I. Ahmad, X. Wang, M. Zhu et al., “EEG-based epileptic seizure detection via machine/deep learning approaches: A Systematic Review,” *Computational Intelligence and Neuroscience*, vol. 2022, no. 311, 2022.
- [38] A. Raza, H. Ayub, J. A. Khan et al., “A hybrid deep learning-based approach for brain tumor classification,” *Electronics*, vol. 11, no. 7, p. 1146, 2022.
- [39] A. Kumar, A. K. Singh, I. Ahmad et al., “A novel decentralized blockchain architecture for the preservation of privacy and data security against cyberattacks in healthcare,” *Sensors*, vol. 22, no. 15, pp. 1–14, 2022.
- [40] N. Muhammad Hussain, A. U. Rehman, M. T. B. Othman, J. Zafar, H. Zafar, and H. Hamam, “Accessing Artificial Intelligence for Fetus health status using hybrid deep learning algorithm (AlexNet-SVM) on cardiocotographic data,” *Sensors*, vol. 22, no. 14, pp. 1–14, 2022.
- [41] A. B. Tufail, N. Anwar, M. T. B. Othman et al., “Early- stage Alzheimer’s disease categorization using PET neuroimaging modality and convolutional neural networks in the 2D and 3D domains,” *Sensors*, vol. 22, no. 12, pp. 1–18, 2022.
- [42] M. S. Mazhar, Y. Saleem, A. Almogren et al., “Forensic analysis on Internet of things (IoT) device using machine-to-machine (M2M) framework,” *Electronics*, vol. 11, no. 7, p. 1126, 2022.
- [43] M. Ahmad, T. M. Ghazal, and N. Aziz, “A survey on animal identification techniques past and present,” *International Journal of Computational and Innovative Sciences*, vol. 1, no. 2, pp. 1–7, 2022.
- [44] S. Muneer and M. A. Rasool, “Systematic review: explainable Artificial Intelligence (XAI) based disease prediction,” *International Journal of Advanced Sciences and Computing*, vol. 1, no. 1, pp. 1–6, 2022.
- [45] A. U. Rahman, R. N. Asif, K. Sultan et al., “ECG classification for detecting ECG arrhythmia empowered with deep learning approaches,” *Computational Intelligence and Neuroscience*, vol. 2022, pp. 1–12, 2022.
- [46] A. M. Shaker, M. Tantawi, H. A. Shedeed, and M. F. Tolba, “Generalization of convolutional neural networks for ECG classification using generative adversarial networks,” *IEEE Access*, vol. 8, pp. 35592–35605, 2020.
- [47] “Kaggle Datasets,” <https://www.kaggle.com/datasets/morriswongch/kaggle-datasets>.

## Research Article

# Understanding the User-Generated Geographic Information by Utilizing Big Data Analytics for Health Care

Hidayat Ullah <sup>1</sup>, Alaa Ali Hameed <sup>2</sup>, Sanam Shahla Rizvi <sup>3</sup>, Akhtar Jamil <sup>4</sup>,  
and Se Jin Kwon <sup>5</sup>

<sup>1</sup>Faculty of Engineering and Natural Sciences, Department of Computer Engineering, Istanbul Sabahattin Zaim University, Istanbul, Turkey

<sup>2</sup>Department of Computer Engineering, Istinye University, Istanbul, Turkey

<sup>3</sup>Raptor Interactive (Pty) Ltd, Eco Boulevard, Witch Hazel Ave, Centurion 0157, South Africa

<sup>4</sup>Department of Computer Science, FAST School of Computing, National University of Computer and Engineering Sciences, Islamabad, Pakistan

<sup>5</sup>Department of AI Software, Kangwon National University, Samcheok 25913, Republic of Korea

Correspondence should be addressed to Se Jin Kwon; [sjkwon@kangwon.ac.kr](mailto:sjkwon@kangwon.ac.kr)

Received 29 July 2022; Revised 8 September 2022; Accepted 17 September 2022; Published 6 October 2022

Academic Editor: Ateeq Ur Rehman

Copyright © 2022 Hidayat Ullah et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

There are two main ways to achieve an active lifestyle, the first is to make an effort to exercise and second is to have the activity as part of your daily routine. The study's major purpose is to examine the influence of various kinds of physical engagements on density dispersion of participants in Shanghai, China, and even prototype check-in data from a Location-Based Social Network (LBSN) utilizing a mix of spatial, temporal, and visualization methodologies. This paper evaluates Weibo used for big data evaluation and its dependability in some types rather than physically collected proofs by investigating the relationship between time, class, place, frequency, and place of check-in built on geographic features and related consequences. Kernel density estimation has been used for geographical assessment. Physical activities and frequency allocation are formed as a result of hour-to-day consumption habits. Our observations are based on customer check-in activities in physical venues such as gyms, parks, and playing fields, the prevalence of check-ins, peak times for visiting fun parks, and gender disparities, and we applied relative difference formulation to reveal the gender difference in a much better way. The purpose of this research is to investigate the influence of physical activity and health-related standard of living on well-being in a selection of Shanghai inhabitants.

## 1. Introduction

Location-Based Social Media Networks (LBSNs) have come a long way since we first took a leap in 2007 when Facebook opened up for use. The LBSN model is one that has been widely used for a variety of purposes including commercial, governmental, and nonprofit works. Though location-based social networking has still been regarded as an innovation, the paradigm is rapidly expanding. People have been taking advantage of the location-tracking data and the resulting insights to build more effective social applications for marketing, advertising, commerce, and even to help in the discovery of real-world phenomena. To date, over one billion

people around the globe have joined social networks, and many of these social networks do not use geocoding to get their information. Instead, they often rely on manually curated lists of user-submitted location identifiers. As a result, a wealth of new information is being generated each day from such a model, and researchers are beginning to mine these records in a systematic way.

The input is typically supplemented with facts, visuals, geo-locations, and textual data, which would be used to make more research on many aspects of people's actions. Past studies used either actually acquired fact for groups in specific classes, including such leisure or LBSN data for the entire society with no preset deployments. If properly

classified, the numerous aspects of the LBSN facts would prove to be a powerful resource of information for assessing people's activity in a variety of fields like as entertainment, education, tourism, dining, and aviation.

As a result, in this research, we will fill the gap in research of employing LBSN statistics in article or content by assessing which amusement places Shanghai citizens desire to attend. Several studies have been conducted in an attempt to analyze and imitate human actions through geo-data. The latest research, for instance, leverages check-in records from globally prominent LBSNs such as Twitter, Facebook, and Foursquare to reveal connections and trends among consumers such as gender, expert or less skilled groups, and age groupings [1–3].

Based on the check-in time, the user may have spent some time at a certain place. This is a potential feature of spatial and temporal aspect. Some spatio-temporal features are related to the check-in time, longitude, latitude, location difference, and check-in time difference. In these features, we can only retrieve information for a location, and the time difference is very small. If we set the threshold of the check-in time difference, then we can select check-in data of different timestamps in the same location. The check-in time can help us to know where a person spent her/his time, and the check-in time difference can help us to understand the daily movement patterns of each user. However, this information does not help us to understand the time spent at different locations or the movement paths in LBSN.

Other researchers, to our knowledge, did not incorporate this earlier. As per significance, we focused on three unique parts of investigation on Weibo's check-in data from the city of Shanghai for two years, from July 2015 to June 2017, to identify spatio-temporal patterns and inhabitant's predictions utilizing physical engagement sites and density prediction. As a consequence, three key aspects of the assessment are highlighted in the newest study. Our input to the existing study is on these topics:

- (i) Time variations of an hour, a week, and a daytime
- (ii) Data collection and physical action site study
- (iii) Using spatial analysis to model and predict density

Section 2 of this study contains relevant material on big data, LBSNs, and the important significance in a range of sectors and also articles on Weibo, Shanghai, and China. Section 3 summarizes both the dataset and the analytical approach. Section 4 contains the facts and a narrative, while Section 5 contains the study's findings and suggestions.

## 2. Related Work

A big issue that we should consider is the question, where do we find the "Big Data?" Actually, even big data analysis does not exist without facts although this does not mean that big data has to include all the data, and sometimes, some data may be more important. In some areas, particularly social sciences, "Big Data" involves massive open online courses (MOOCs), blogs, wikis, video recordings, live tweets, live chats, microblogs, video images, and documents that capture

data (i.e., digital documents, websites, images, and videos) in a manner that enable the monitoring and management of these data over time and space, and this term is more often associated with Internet data and specifically Internet Big Data although the term "Internet Big Data" can also be applied to traditional data sources too. However, Big Data is not to be considered the opposite of "small data," with the latter referring to the data that is more easily handled and analyzed [4–8]. Instead, it is the integration of and the interrelation between small and large datasets.

Big Data is a type of data, collected and processed in a structured and unstructured way. On the one hand, a Big Data application is focused on collecting data that is big, complex, and diverse. On the other hand, another Big Data application has a focus on transforming the collected data into information and knowledge. Ovadia demonstrated and emphasized the importance of "Big Data" for intellectuals and social scientists, claiming that "Big Data" is too important to overlook since many social scientific study needs a large amount of facts and enormous datasets [9].

Several research fields, including time and human mobility, space geography, human behavior, and metropolitan activities, originated with statistical data collected via surveys, visit records, questionnaires, interviews, and other hand-crafted datasets [10, 11]. These methodologies comprising geo-information, have recently spread powerful resources of information for such scientific papers, [6, 12, 13]. Tracking users' movements and actions have become simple because of the fast progress of mobile technology and the extensive use of mobile devices [14]. Though giving an approximated space just next to the mobile's base transmitter in which the calls were routed, the dataset demonstrated competence in forecasting user positions with little time, which was then used to anticipate user activities [15].

Zhu presented numerous GIS (geographic information system) components and their importance in pattern extraction and municipal studies by demonstrating how to evaluate and display the spatio-temporal parts of reusable garbage, collection, and rehabilitation [16–18]. The authors in these researches utilized these data for health purposes [19–25] and medical data security [18, 26, 27]. These researches are based upon endoscope image, medical image registration, and soft tissue modeling [28–30].

The term "business intelligence" (BI) first appeared in the late 1990s [31]. At the organizational level, "Big Data" is a significant area. Prediction, ad hoc inquiries, and aggregation-based reporting, along with processing organized and unorganized data and combining "Big Data"-based systems, all contribute to better decision-making [32–34]. The BI systems include data warehouses for gathering clean, precise, and comprehensive data from a variety of origins, and also online analytical processing (OLAP) for real-time multidimensional analysis with processes, for example, combining, screening, roll-up, spin, and exploration for details [35].

OLAP is such a well-recognized and well-respected methodology for "Big Data" research in BI systems [36]. However, while BI, OLAP, and data warehouse are powerful tools for coping with vast volumes of data and a vast variety of operations, they also provide a challenge owing to the

considerable cost, space, and computing resources necessary [37–40]. The authors introduced new data mining techniques in these researches [41–43]. Digital social networks have been demonstrated to be the most significant source of “Big Data” to research personal activity since they have been used and are fast expanding in virtually every location across the world.

Customers of the LBSNs’ digital services may upload and exchange their actions, preferences, and whereabouts, resulting in massive amounts of data for numerous researches on many areas. Papers [44, 45] go into excessive details regarding the human behavior analytic techniques. Lindqvist [46] discovered the usage of LBSNs, which was encouraged by earlier research studies; for example, practical studies and socio-spatial aspects utilizing LBSNs [47] and a personalized geo-social suggestion relied on a dataset from two independent LBSNs, namely, Gowalla and Foursquare [48,49]. By gathering regular users at various locations, researchers in two UK cities used comparable check-in data from LBSN to enhance the recommendation algorithms.

Li [50] did a thorough examination of location-based data from 2.4 million places in 14 states to determine the elements affecting place popularity. The research’s findings identified three major factors that influence a location’s fame: site profile, site age, and site type. Another study on consumer actions at numerous venue categories focused on “Food” in Riyadh, Saudi Arabia, and revealed that when clients attend food venues, they seem to be more receptive to discussing their behaviors. The check-ins of about 19,000 Swarm (Foursquare) members from three metropolitan centers, mainly San Francisco, Hong Kong, and New York were utilized to evaluate linkages between different sites at different times of the day [51].

Several researches have been conducted all across the universe to investigate different consumer and check-in characteristics utilizing LBSN sources of information like Twitter and Foursquare. These traits have been used in a variety of sectors, together with transportation trends, venue categorization, and urban development and growth [52]. Weibo, a well-known Chinese LBSN, was used in this investigation and demonstrated to be beneficial. In research for Shenzhen [53], Weibo data were utilized to analyze the request powers of tourist charms. Another study on people’s action trends and engagements was assumed in order to observe urban boundaries in Beijing [54] which likewise used Weibo check-in. In a comparable vein, Shi et al. [52] utilized Weibo data to examine attributes of tourism sites based on numerous details provided by the LBSN, and the evaluation was combined with feelings from user reviews. Wu et al. [55] performed spatio-temporal assessment grounded on the period of day and the variance in check-in patterns between weekdays and weekends [56, 57] also examined check-in research containing 21 of Wuhan’s most famous lakes.

### 3. Dataset and Methodology

*3.1. Study Area and Data Source.* Shanghai, like the other megacities in East Asia, experiences urban sprawl. This is particularly the case in low- and middle-income groups and

among minority groups, such as the ethnic minorities (like Zhuang, Dong, Miao, and Hui), while the rich and high-income groups have higher purchasing power are more likely to move into the inner city and tend to be concentrated in the central city. The hour of the day and the difference in check-in trends across weekdays and weekends were used to do a spatio-temporal evaluation. Figure 1 displays the research domain, and it can be seen that we focused and utilized the data of 10 districts.

The data are gathered from Weibo, which contains people’s check-in details while they are at a specific place. The check-ins from Weibo contain latitude, longitude, accuracy (e.g., 1.8 meters), address, time, and content. In addition, Weibo has been updated more than 500 million times in a day, covering about 1/5 of all of Weibo’s activity, which makes it one of the most extensive and authoritative datasets in the field of big data. The spatio-temporal check-ins gathered in this research is collected from the city of Shanghai, which covers four prefectures. There are three major problems associated with collecting the spatio-temporal check-ins of the Weibo from a single city. First, it is very difficult to extract real locations from a large number of posts with various locations. Thus, we have to use heuristics such as the latitude and longitude of the home location to extract places. Second, as Weibo user accounts contain a location-based tag, the real location of the user is more or less accurate, but the real location of posts is different from the location of the corresponding Weibo users. Third, some posts just contain a location without other information such as a picture. The number of visits has risen to 500 million by the end of 2018, with 462 million monthly regular individuals and 200 million daily active users. This study focuses on two years of socially collected spatio-temporal check-ins from Weibo in Shanghai, from July 2015 to June 2017.

The fundamental reason for utilizing LBSN is to exchange activities and observations, which results in the establishment of a new close social fellowship group. This allows experts to deduce a broad spectrum of individual action and pleasure from the geo-data collected by these LBSNs. This study’s data originated from LBSN’s Weibo account. We utilized the Weibo API (application programming interface) based on Python to collect data from check-ins in Shanghai.

This was compiled in 2017 within over 3.5 million cumulative check-ins from approximately two million people. For the current study, the data were converted from JavaScript Object Notation (JSON), the standard API Java programming language, to comma-separated values (CSV) using Mongo DB. The information processing route is depicted in Figure 2.

JSON is a tiny data transfer standard that uses human-readable language to transmit data entities, while Java is an object-oriented programming environment [58]. The information was combined into one file in the CSV (comma-separated values) format for more processing and analysis with the tools indicated.

All of the contributors’ information, including geo-location, might be recorded in a database. We collected the data in the CSV format and then used a criterion to

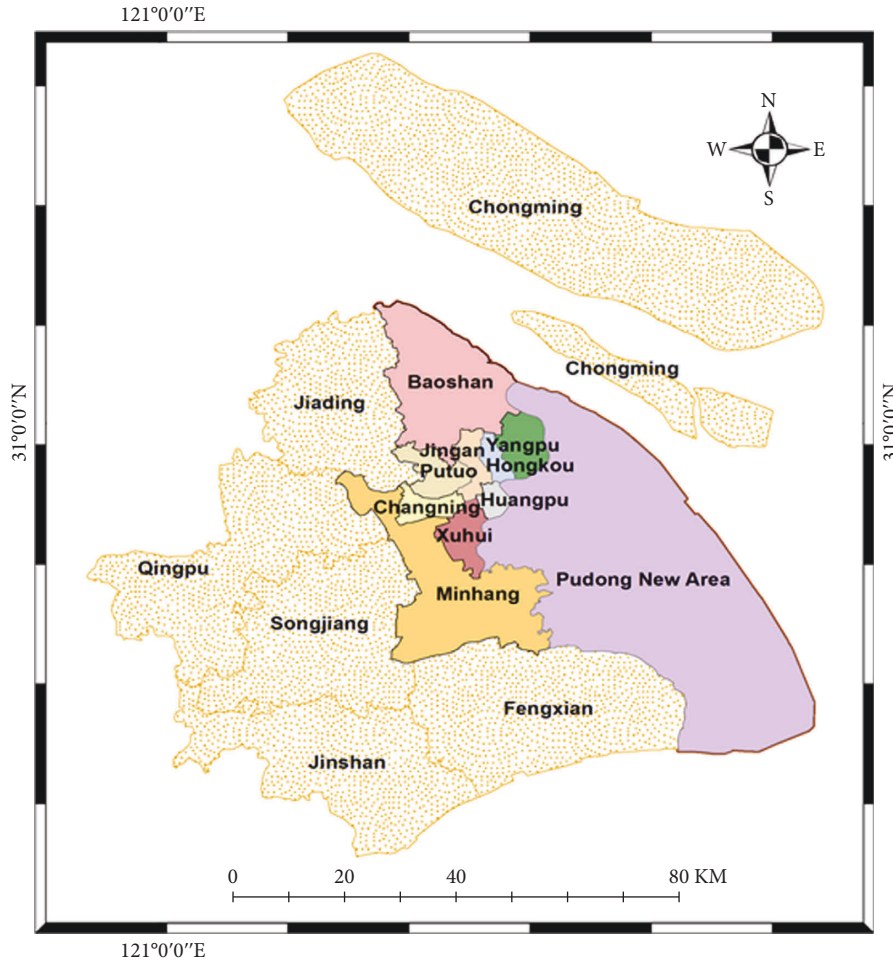


FIGURE 1: Study area.

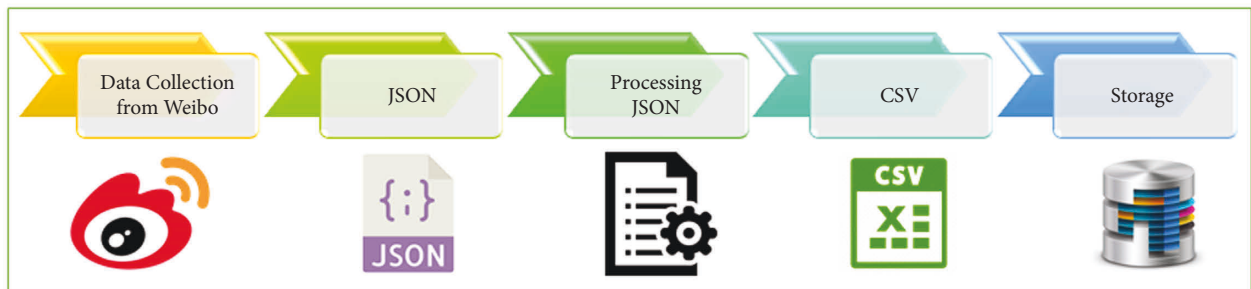


FIGURE 2: Process flow of data.

determine the significance of the results. The criterion figure is depicted in Figure 3.

The CSV standard is the most prevalent and significant standard for databases and spreadsheets, and it uses commas as delimiters for different parameters. In the JSON file standard, keys (ID, latitude, longitude, etc.) are used as headers for the CSV file format, while values (5404478798, 121.544449, 31.268159, etc.) are used as descriptive data. In the CSV file, Table 1 shows an illustration of a “check-in.”

3.2. Social Media Data Analytics Framework. A social media data analytics framework was constructed to provide the

analytical capability of a business unit or an enterprise organization that needs to use the social media to analyze the social media data in their systems. Social media data are the social media messages, tweets, blog posts, pictures, and so on created by the user(s). Social media data analytics framework can analyze this data, process them, and provide the required information. Social media data analytics tools are tools that perform real-time data analytics and report on social media. They are used to analyze the social media data, discover valuable insights about social data, and deliver actionable recommendations. Our broad geographical assessment approach is depicted in Figure 4. The first element is broken into two parts: data collection (downloading Weibo data)

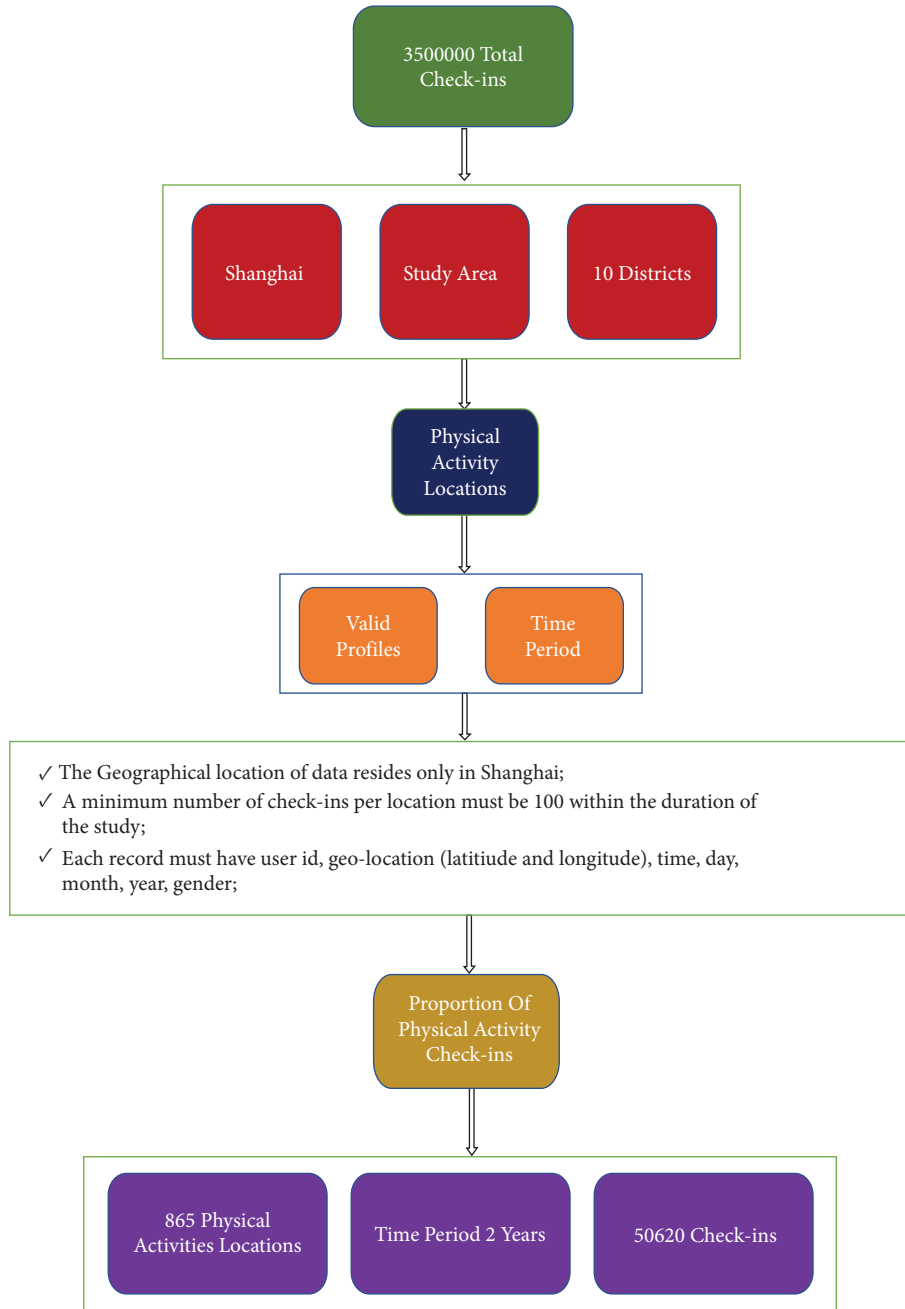


FIGURE 3: Criteria of research.

TABLE 1: Example of Weibo check-in.

Building_id	User_id	Date	day	Time	Year	Gender	Lon	Lat	Address
2930091850	###*	24	Wed	0:00:03	2016	F	121.58689	31.2108	Sports and leisures
3785887251	###*	11	Fri	1:24:45	2016	F	121.3710578	31.14217173	Sports and leisures
3949989322	###*	29	Mon	3:02:19	2017	M	121.44087	31.18104	Sports and leisures

and data cleaning. Following that, the LBSN data is analyzed , as shown in Figure 5.

We used ArcGIS software for the spatial analysis and tableau for the graphical representation of temporal analysis.

3.3. *Spatial Method.* KDE is used to estimate the distribution of data and can be useful in estimating probability distributions of numerical variables and to calculate a density-based visual perception in a given image

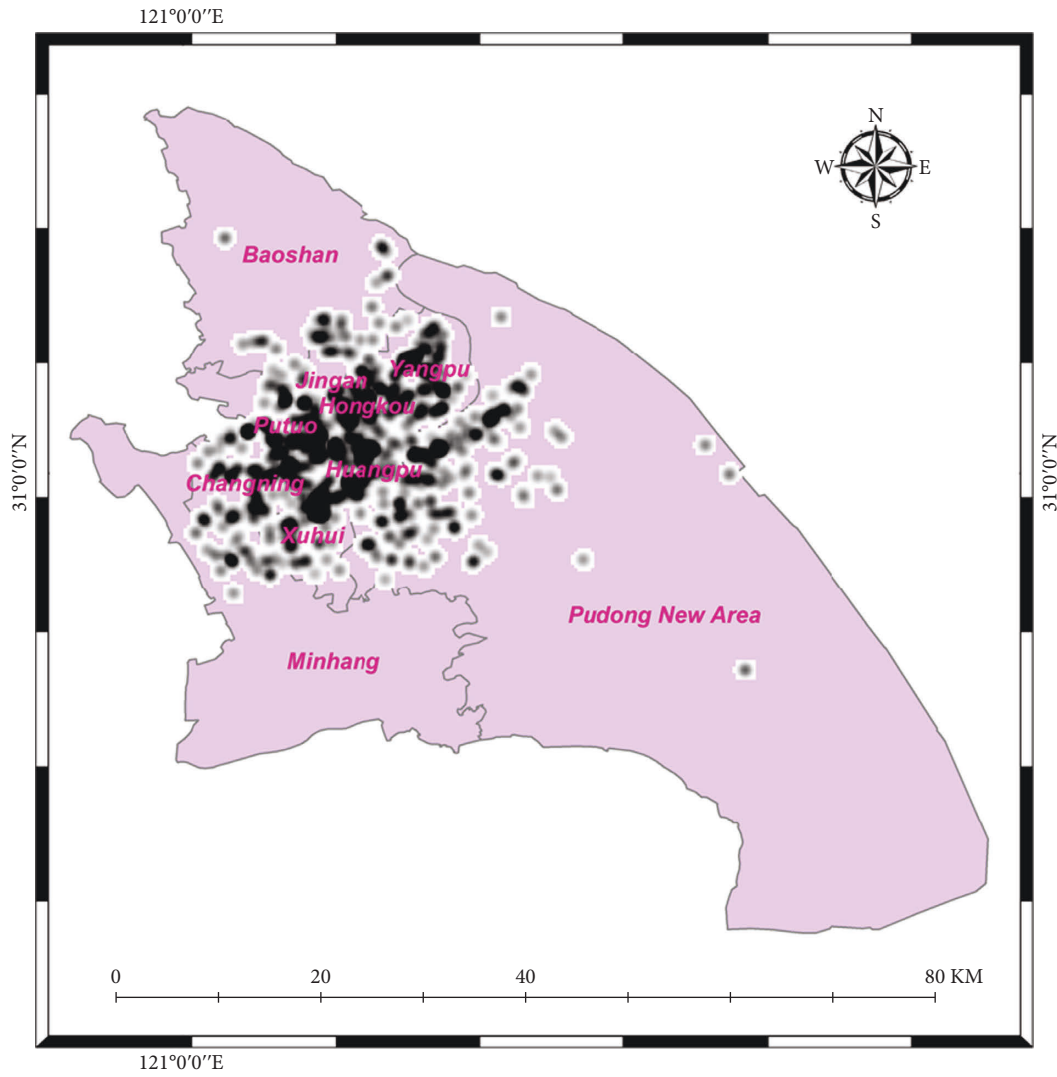


FIGURE 4: Check-in density of activities.

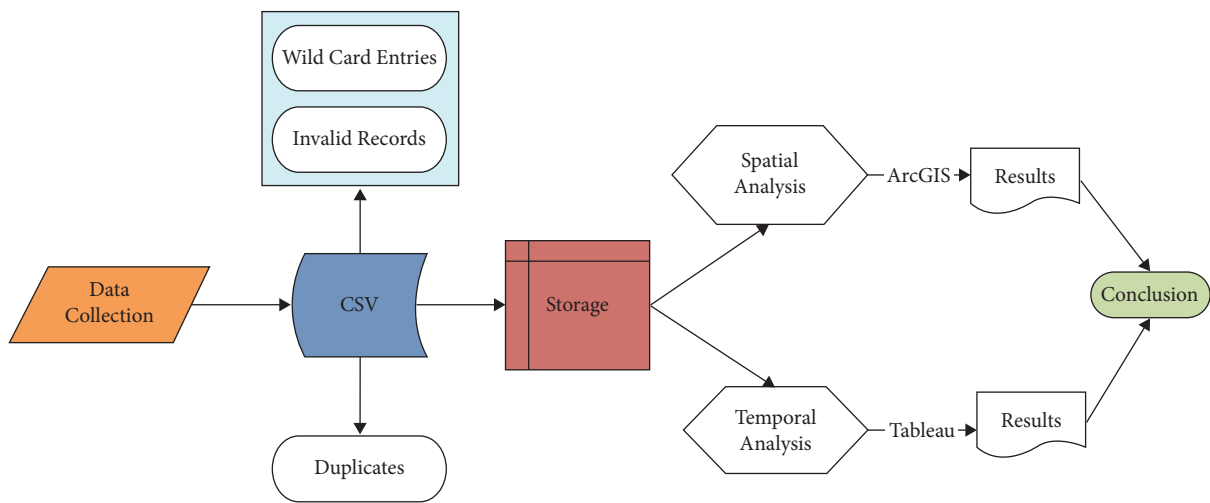


FIGURE 5: Research methodology.

[47, 48, 52, 56, 59–68]. The method works by determining an exact probability distribution function in a given image and then calculating the integral of this probability distribution function to obtain corresponding intensity estimation. There are a number of different estimation techniques for numerical variables. The most frequently used methods are the kernel density estimate (KDE), the Parzen window (PW), and the Epanechnikov window. As the name indicates, Kernel density estimate uses the kernel function as its basis for a density estimate. The Parzen window method is a variation of KDE that uses a Gaussian kernel function. The Epanechnikov window technique is a nonparametric density estimation approach that makes use of the Epanechnikov window's flattening capabilities. Since the 1960s, the Epanechnikov window has been utilized as a nonparametric concentration prediction approach. However, the KDE and PW methods are parametric methods. The Epanechnikov window method is a nonparametric method. Parametric methods are based on distribution functions. To compute a parameter estimate using a parametric method, it is necessary to specify the parameter values. Parametric methods are also commonly used in image analysis and computer vision. A parametric method assumes that the object that is to be described has a unique parameter. Parametric density estimation is widely used to estimate the parameters in the distribution of physical quantities and to estimate probability distributions.

We utilized the KDE method here in our research because for modeling spatial-densities, the KDE method has also been used in fields such as health, marketing, and environment [48, 60, 61]. Recently, as it is used in the analysis of spatial data for determining distribution of the phenomena in geospatial data, many researchers applied the technique to this field. One of the basic topics is that the density of points is determined, but the points are distributed on grids due to the data structure. There are many methods of determining the density. If the grid size is large, many points have been concentrated in each grid, and the points are not sparsely distributed. Thus, the density of the points is high. Thus, it is thought that the density is determined by using the density of the grid. This process is called density-grid transformation.

KDE has been used in studying the patterns of visitors in green parks [66, 69–77]. These studies, mostly concerned with the visitors' number, are based on the assumption that people who visit a green park have similar behavior. To identify the number of people who visit a green park, the visitors are often modeled using kernel density estimation. The number of visitors in this method is simply defined by the area of each kernel. The mean of these individuals is called the mean estimator of the total number of people who visit a green park. It is not known whether this method gives an unbiased estimate of the number of people who visit a green park.

Let  $E$  be a set of historical data where  $e^j = \langle x, y \rangle$  is the geo-coordinates of a location,  $1 \leq j \leq n$  for an individual  $i$ ,  $h_j$  is the Euclidean distance to  $k$ -th near neighbor  $e^j$  in the training data. The KDE is stated as follows:

$$f_{KD}(e|E) = \frac{1}{n} \sum_{j=1}^n K_{h_j}(e, e^j), \quad (1)$$

$$K_h = \frac{1}{2\pi h} \exp\left(-\frac{1}{2}(e, e^i)^t \sum_h^{-1}(e, e^i)\right).$$

#### 4. Results

Shanghai City, with such a population of 22,125,000 residents and a land area of 4,015 square kilometers, has become one of the world's fastest-growing cities [78, 79]. Over two years, data have been accumulated on amusement check-ins. Every single check-in was allotted a value that best suited the physical activities taken out at that place, like gym, sports, and park exercise. Figure 6 depict the total locations of such activities in our study area, and the total number of locations is 865, as can be perceived in the given figure.

We investigated the geographical variation of check-in data with KDE and showed the Weibo geolocation check-in dataset with ArcGIS. From July 2015 to June 2017, the overall check-in intensity in Shanghai was depicted in Figure 4. Sections highlighted in black represent a bigger number of persons, a higher frequency of activity, and a better knowledge of social network usage. It is just not surprising that the seven districts look thicker than the other three districts even though the three districts have a bigger area.

Figure 7 demonstrates the temporal fluctuations in the number of visitors over 24 hours. Even though visitors participated at all hours of the day, most check-ins were recorded between 05 PM and 09 PM among the entertainment sites investigated. Until midnight, the tendency will continue to rise.

Figure 8 illustrates the digit of check-ins for each day with the gender difference, and it can be observed that weekends have a bigger size of check-ins than weekdays but the number of males is higher than that of females for each day, and it is shocking.

Check-ins are disparaged at the district level to provide a more accurate picture of entertainment place distribution in Shanghai City. Figure 9 shows that the distribution of check-ins is the greatest in the Pudong region, preceded by the Huangpu district.

This pattern can be explained by the fact that the Pudong new area district is larger than some other districts. Another factor to bear in mind is that check-in dispersion is stronger in the city region than in the outskirts. Gender differences may also be evident, with male check-ins outnumbering female check-ins throughout all areas.

Figure 10 illustrates the total overall check-ins made each day in all the districts of study area. It can be witnessed that if it is the day or a district, the number of check-ins made by men is higher than that of females.

We applied an analysis on our given data for revealing the comparison of gender difference in the physical strain activities which are really necessary for a healthy life. Table 2 shows the percentage of number of check-ins.



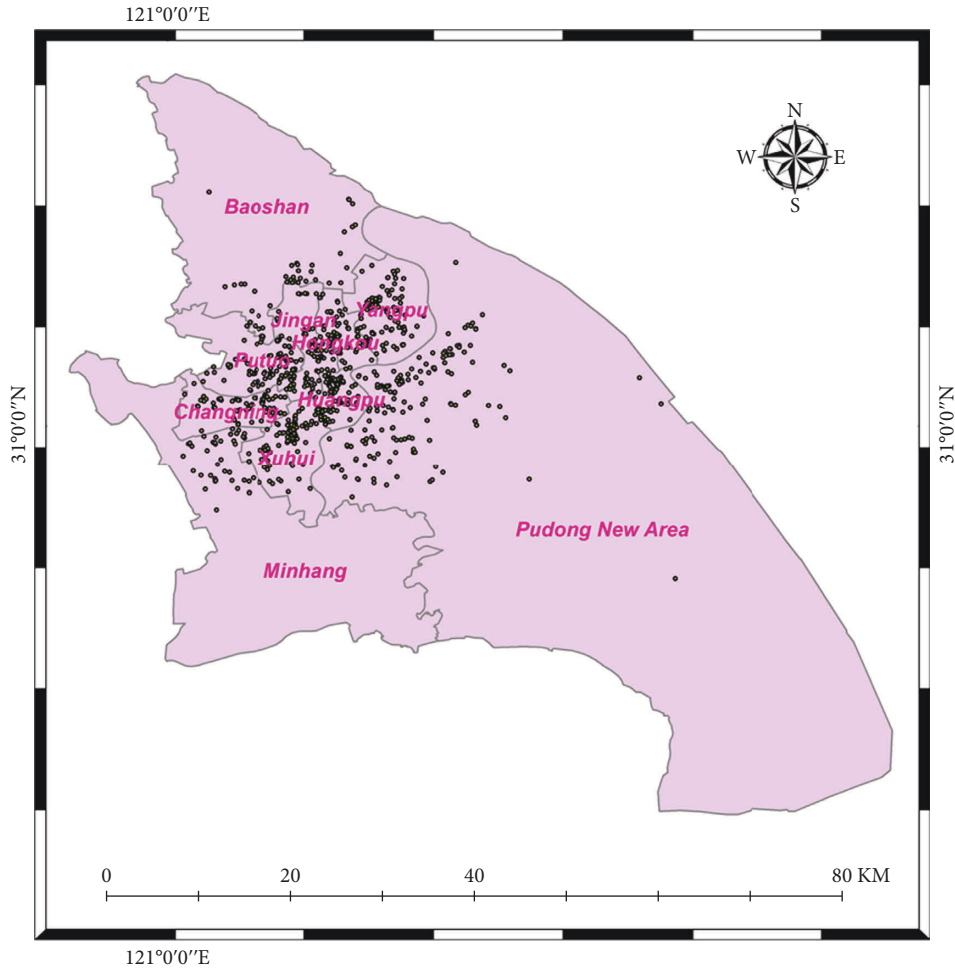


FIGURE 6: Locations of activities.

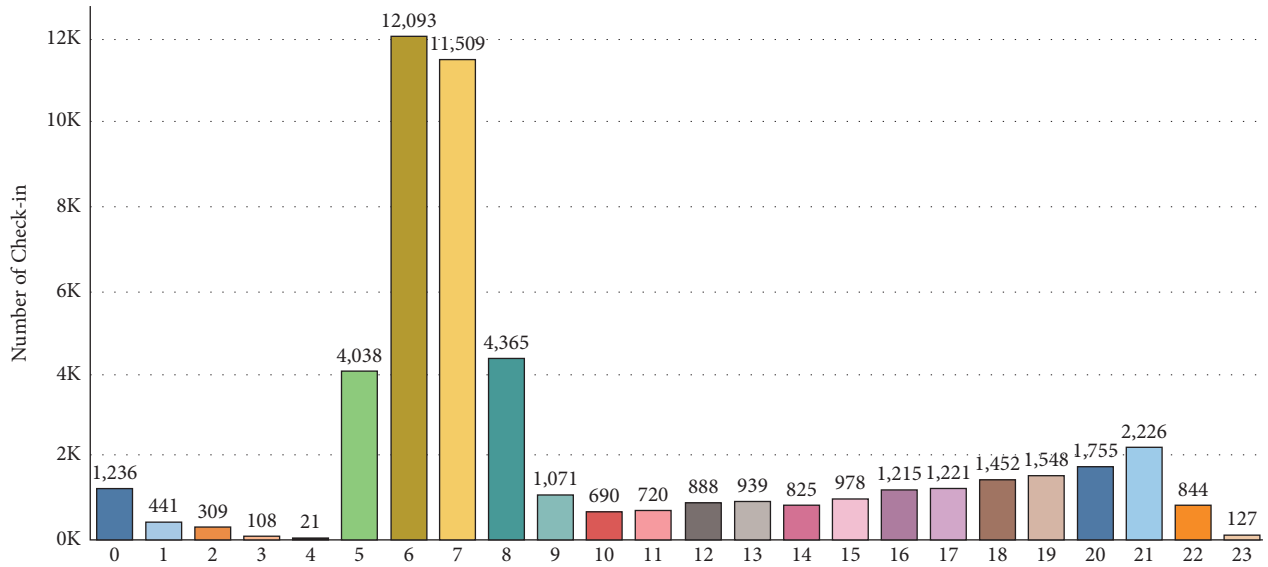


FIGURE 7: Hourly check-in frequency.

To compare districts and days, we used a relative difference ( $d_r$ ) of a given gender. When there is a big difference, the absolute difference is small. If the difference

between females and males is large, the relative difference is smaller. For example, we calculate the difference between total male check-ins and male check-ins per day in

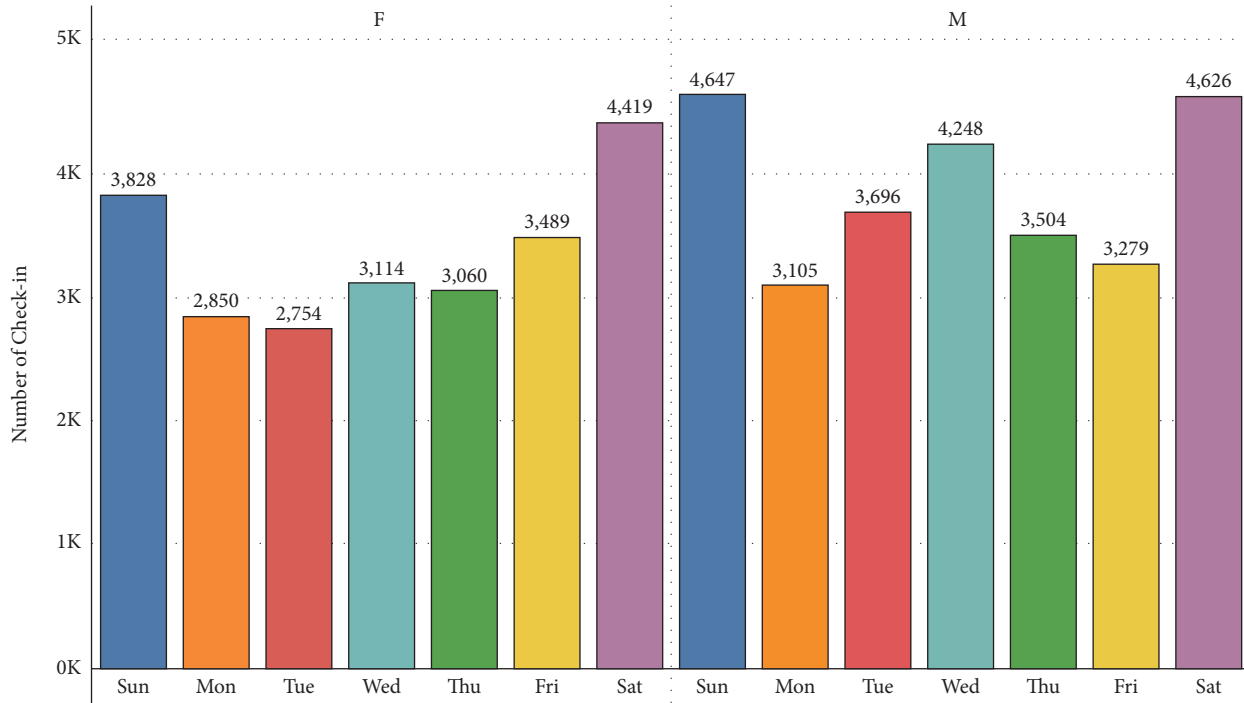


FIGURE 8: Daily number of check-in and gender difference.

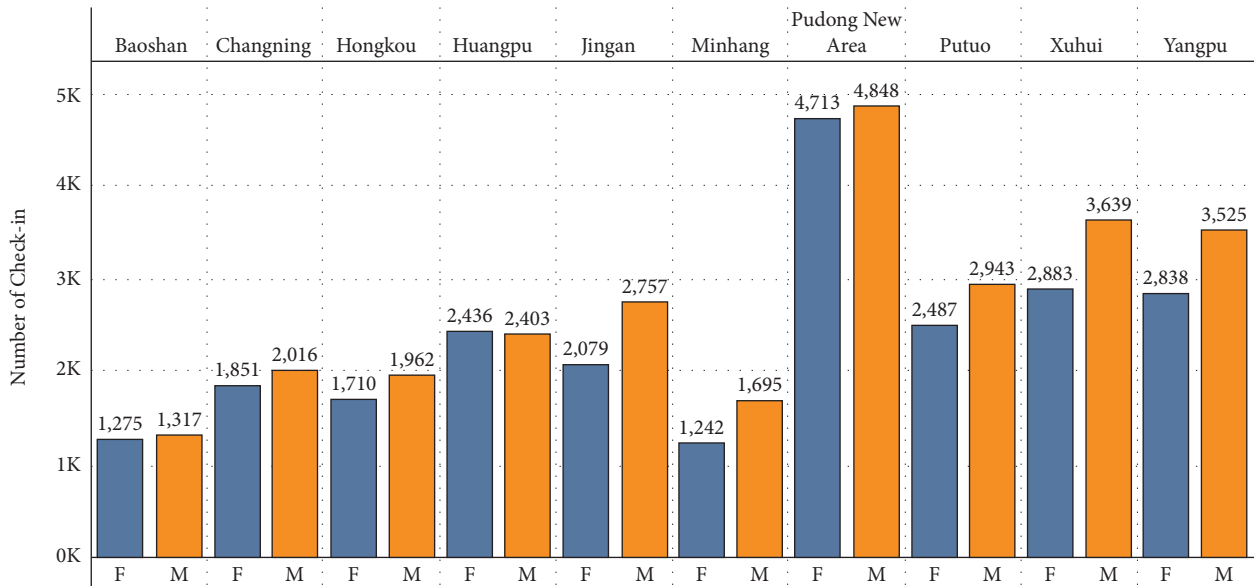


FIGURE 9: Check-ins distribution in districts.

weekdays in the districts. There are some districts with a relatively large difference, and there are some districts where the difference is small. It is generally used as a quantitative pointer of quality control and quality assurance and is stated as follows:

$$d_r = \frac{|P_m - P_f|}{(|P_m| + |P_f|/2)} \quad (2)$$

Finally, the analysis examines the difference between all-male group and female-male group in Shanghai and the time span of two years to form male-only groups and female-only groups. In this context, all-male group means that the total check-ins of male is more than that of females, whereas the female-only group means that females' check-in is more than males' check-in. Tables 3 and 4 shows that there are significant differences among gender in both days in week and districts. The results show that the frequency of women

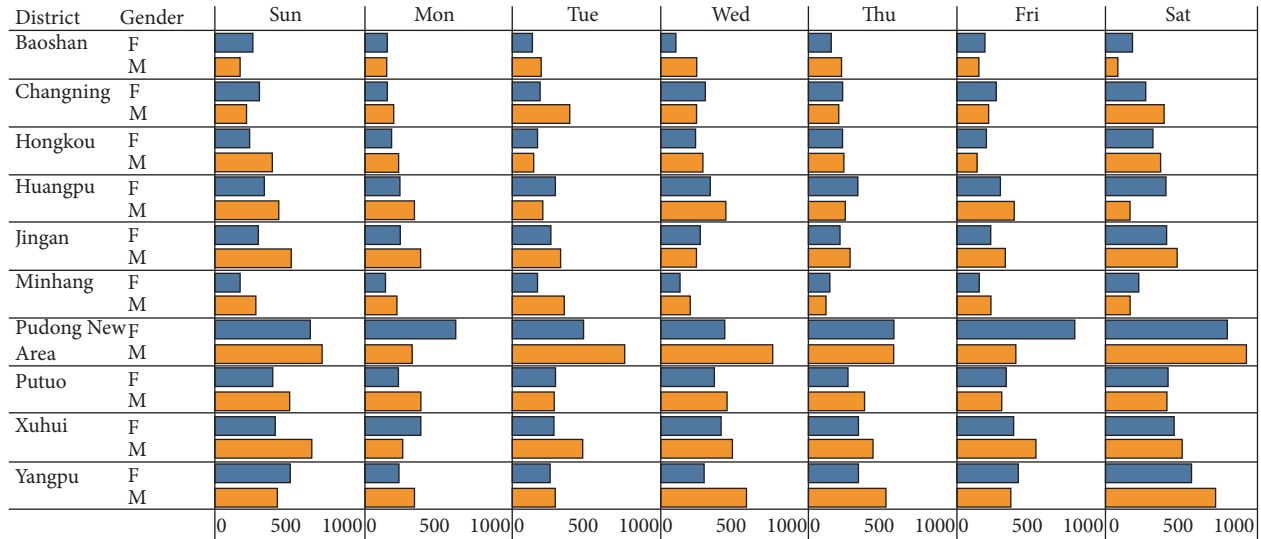


FIGURE 10: Overall statistics.

TABLE 2: Overall statistics.

Gender	District	Sun (%)	Mon (%)	Tue (%)	Wed (%)	Thu (%)	Fri (%)	Sat (%)	
F	Baoshan	0.55	0.33	0.30	0.23	0.33	0.40	0.39	
	Changning	0.63	0.34	0.41	0.64	0.49	0.56	0.57	
	Hongkou	0.50	0.41	0.37	0.50	0.49	0.43	0.68	
	Huangpu	0.72	0.53	0.62	0.73	0.72	0.62	0.87	
	Jingan	0.62	0.54	0.55	0.57	0.46	0.47	0.88	
	Minhang	0.37	0.31	0.37	0.28	0.32	0.32	0.48	
	Pudong New Area	1.37	1.32	1.03	0.92	1.23	1.69	1.74	
	Putuo	0.84	0.50	0.63	0.78	0.57	0.69	0.90	
	Xuhui	0.87	0.83	0.60	0.87	0.72	0.81	0.99	
	Yangpu	1.08	0.52	0.56	0.62	0.72	0.89	1.22	
	M	Baoshan	0.37	0.33	0.42	0.51	0.49	0.31	0.18
		Changning	0.46	0.44	0.83	0.52	0.44	0.45	0.84
Hongkou		0.83	0.52	0.31	0.62	0.51	0.30	0.79	
Huangpu		0.92	0.74	0.44	0.94	0.53	0.82	0.35	
Jingan		1.10	0.80	0.70	0.52	0.60	0.69	1.04	
Minhang		0.58	0.48	0.75	0.43	0.25	0.50	0.35	
Pudong new area		1.54	0.69	1.61	1.62	1.24	0.85	2.03	
Putuo		1.08	0.83	0.60	0.97	0.80	0.65	0.88	
Xuhui		1.39	0.57	1.01	1.04	0.94	1.14	1.10	
Yangpu		0.89	0.73	0.63	1.23	1.13	0.78	1.58	

in the districts is less than the men in all days of the week, except Friday. This indicates that male users prefer other days to Fridays while women users are more active on Friday. Table 3 shows the difference between male and female user check-ins on the different days of the week. As expected, we see significant difference among the days of the week. Overall, Saturday has the highest number of male check-ins, and Friday has the highest number of female check-ins. This indicates that different days of the week are equally important in terms of check-ins. However, we see some differences in the number of check-ins among the different districts.

We see some differences in the number of check-ins in the different districts. Table 4 shows that districts Jingan, Putuo, Xuhui, and Yangpu have a significant difference between male and female check-ins. These results confirm

the previous observation that there is difference in the distribution of the male and female users in the city.

## 5. Discussion

The study of urban environment is attracting increasing research interest and has developed rapidly in recent years. However, there is still a gap between academic and popular knowledge about activities, and most existing research are conducted in the form of case studies in one or two places. Most of the research data and conclusions are based on the investigation of one city or district. The research results have limited significance for the whole city. However, most study data are examined using statistical approaches that do not take into account spatio-temporal features, so they cannot represent the popularity of activity across time and location.

TABLE 3: Gender Differences during week.

Day	Check-in (%)	Female (%)	Male (%)	$d_r$
Sun	16.74	7.56	9.18	0.193
Mon	11.76	5.63	6.13	0.085
Tue	12.74	5.44	7.30	0.291
Wed	14.54	6.15	8.39	0.308
Thu	12.98	6.05	6.92	0.135
Fri	13.37	6.89	6.48	0.062
Sat	17.87	8.73	9.14	0.046

TABLE 4: Gender difference in districts.

Day	Check-in (%)	Female (%)	Male (%)	$d_r$
Baoshan	5.12	2.52	2.60	0.032
Changning	7.64	3.66	3.98	0.085
Hongkou	7.26	3.38	3.88	0.137
Huangpu	9.56	4.81	4.75	0.136
Jingan	9.56	4.11	5.45	0.281
Minhang	5.80	2.45	3.35	0.309
Pudong new area	18.88	9.31	9.58	0.028
Putuo	10.72	4.91	5.81	0.168
Xuhui	12.89	5.70	7.19	0.232
Yangpu	12.57	5.61	6.96	0.216

This not only contradicts the features of the urban area but also fails to suit the demands of modern city design. In this paper, we aim to analyze the characteristics of the popular physical activities in Shanghai from a spatio-temporal perspective and compared the gender difference. We study how the Shanghai public’s popularity for different kinds of physical activities has changed over time, which popular location is most favored by the public, and in which places the physical activity is most favored. The data of Shanghai Weibo, the largest social media platform in China, is used to conduct the investigation. In summary, we find those as follows: (1) There are some places with higher popularity compared to others and (2) there is also difference in the popularity of physical activities over time and space. We believe that these findings can not only provide references for urban management but also help to understand the public’s favorite locations in one city.

Several limitations exist in this study. Data security is really necessary for dataset or any method [80, 81]. First, since the data we collected were based on online records, the results may be biased because different Internet access habits among different geographical areas affect the data. For example, people in rural areas may have more Internet access and thus have more chances of accessing social media than people in urban areas. Second, we only selected a single city as the sample, and this limited the generalization of our results. Third, although WeChat check-in data is a very powerful tool in the evaluation of physical activities quality, this approach still contains subjective information, which cannot be quantified. Therefore, the evaluation of data quality in our study should be revised based on more and more data-driven evaluation methods.

Accessibility of the data is a major hurdle to LBSN research due to data and confidentiality safety considerations. The ability of LBSNs to reveal users’ and their contacts’ present geolocation which raises serious privacy concerns. They are worried about their privacy, so are administrative or business users that communicate data over LBSNs. Personal information is sometimes given freely or inadvertently. Though the data are acquired frequently by offering customers special rights and prizes in exchange for their information, this is not at all times the case.

Our findings indicate that this new channel has been relatively successful in terms of both social media check-in and activity tracking activity. However, despite this success, this channel only covers roughly half of all locations. For this reason, our findings have important implications for how location-based data might be used. In terms of data validity, the study provides a unique opportunity to compare Weibo check-ins with activity tracking. Only a few studies have compared the two channels. As with other research, we found that the Weibo data were accurate, but not complete. As discussed in the introduction section, we believe this is not a reflection of the data validity but rather an aspect of the usage of these services.

Our findings contribute to the understanding of user behavior and the understanding of the behavior of social media users in general. Because check-ins are executed by a moderately small group of Weibo users, they have been used in other studies as a good pointer of Weibo users’ day-to-day activity [71]. However, our findings reveal that the number of locations has not been increased significantly in Shanghai. When compared to prior studies, it can be noted that all of the studies show that female consumers are much more active than males during the various activities in Shanghai [69–74, 82, 83].

## 6. Conclusion

We looked at consumers’ check-ins in 10 distinct districts of Shanghai, focusing on different aspects of geo-referenced data. We utilized the KDE method by collecting the check-ins of different users of Weibo microblog and all the check-ins were collected during the users posted when they were doing any physical activity in gyms, playing grounds, and other parks as well. Our data shows that most people submitted their physical action check-ins in Shanghai’s city center, which is separated into seven districts. Pudong new area and Huangpu are denser districts than others, and weekends have more check-ins compared with other days. The conclusions reveal that male users are more active in physical activities either in gym or playing ground in all the districts of Shanghai and its shocking revelation. The purpose of our research is to create awareness among people about the benefits of physical activities on health. The research may be valuable in detecting more overcrowded areas in Shanghai so that regulating or management institutions can more efficiently watch and aid such districts, notably in events, public action, and urban development, among other things.

## Data Availability

The dataset can be acquired from the correspondence upon request.

## Conflicts of Interest

The authors declare that there are no conflicts of interest.

## References

- [1] N. Alrumayyan, "Analyzing user behaviors: a study of tips in Foursquare," in *5th International Symposium on Data Mining Applications*, Springer, Germany, 2018.
- [2] J. Preston and B. Stelter, *How Government Officials Are Using Twitter for Hurricane Sandy*, The New York Times, New York, NY, USA, 2012.
- [3] L. Hou, Q. Liu, M. Uddin, H. Khattak, and M. Asshad, "Spatiotemporal analysis of residents in shanghai by utilizing Chinese microblog Weibo data," *Mobile Information Systems*, vol. 10, 2021.
- [4] P. Chatterjee, "Big data: the greater good or invasion of privacy," *The Guardian*, vol. 12, 2013.
- [5] E. Dumbill, "What is big data: an introduction to the big data landscape (article)," *Strata Oreilly*, vol. 11, 2012.
- [6] M. Graham and T. Shelton, "Geography and the future of big data, big data and the future of geography," *Dialogues in Human geography*, vol. 3, no. 3, pp. 255–261, 2013.
- [7] S. Lohr, "Big data is opening doors, but maybe too many," *New York Times*, vol. 23, 2013.
- [8] V. Mayer-Schönberger and K. Cukier, *Big Data: A Revolution that Will Transform How We Live, Work, and Think*, Houghton Mifflin Harcourt, 2013.
- [9] S. Ovadia, "The role of big data in the social sciences," *Behavioral & Social Sciences Librarian*, vol. 32, no. 2, pp. 130–134, 2013.
- [10] Y. Chai, "Review for space-time behavior research: theory frontiers and application in the future," *Progress in Geography*, vol. 31, no. 6, pp. 667–675, 2012.
- [11] Z. Lv, S. Lv, H. Feng, H. Zhu, and H. Lv, "Clinical characteristics and analysis of risk factors for disease progression of COVID-19: a retrospective Cohort Study," *International Journal of Biological Sciences*, vol. 17, pp. 1–7, 2021.
- [12] K. Wang, B. Zhang, F. Alenezi, and S. Li, "Communication-efficient surrogate quantile regression for non-randomly distributed system," *Information Sciences*, vol. 588, pp. 425–441, 2022.
- [13] F. Zhang, J. Zhai, X. Shen, O. Mutlu, and X. Du, "POCLib: a high-performance framework for enabling near orthogonal processing on compression," *IEEE Transactions on Parallel and Distributed Systems*, vol. 33, no. 2, pp. 459–475, 2022.
- [14] M. C. Gonzalez, C. A. Hidalgo, and A.-L. Barabasi, "Understanding individual human mobility patterns," *Nature*, vol. 453, no. 7196, pp. 779–782, 2008.
- [15] C. Song, Z. Qu, N. Blumm, and A. L. Barabasi, "Limits of predictability in human mobility," *Science*, vol. 327, no. 5968, pp. 1018–1021, 2010.
- [16] X. Zhu, "GIS and urban mining," *Resources*, vol. 3, no. 1, pp. 235–247, 2014.
- [17] Z. Cao, Y. Wang, W. Zheng et al., "The algorithm of stereo vision and shape from shading based on endoscope imaging," *Biomedical Signal Processing and Control*, vol. 76, Article ID 103658, 2022.
- [18] M. Zhang, Y. Chen, and J. Lin, "A privacy-preserving optimization of neighborhood-based recommendation for medical-aided diagnosis and treatment," *IEEE Internet of Things Journal*, vol. 8, no. 13, pp. 10830–10842, 2021.
- [19] A. Kumar, A. K. Singh, I. Ahmad et al., "A novel decentralized blockchain architecture for the preservation of privacy and data security against cyberattacks in healthcare," *Sensors*, vol. 22, no. 15, p. 5921, 2022.
- [20] N. Muhammad Hussain, A. U. Rehman, M. T. B. Othman, J. Zafar, H. Zafar, and H. Hamam, "Accessing artificial intelligence for fetus health status using hybrid deep learning algorithm (AlexNet-SVM) on cardiocardiographic data," *Sensors*, vol. 22, no. 14, p. 5103, 2022.
- [21] S. Kanwal, "A robust data hiding reversible technique for improving the security in e-health care system," *CMES-COMPUTER MODELING IN ENGINEERING & SCIENCES*, vol. 134, 2022.
- [22] I. Ahmad, I. Ullah, W. U. Khan et al., "Efficient algorithms for E-healthcare to solve multiobject fuse detection problem," *Journal of Healthcare Engineering*, vol. 2021, Article ID 9500304, 2021.
- [23] A. U. Rehman, R. A. Naqvi, A. Rehman, A. Paul, M. T. Sadiq, and D. Hussain, "A trustworthy snot aware mechanism as an enabler for citizen services in smart cities," *Electronics*, vol. 9, no. 6, p. 918, 2020.
- [24] M. Khushi, K. Shaukat, T. M. Alam et al., "A comparative performance analysis of data resampling methods on imbalance medical data," *IEEE Access*, vol. 9, pp. 109960–109975, 2021.
- [25] K. Shaukat, S. Luo, V. Varadharajan et al., "Performance comparison and current challenges of using machine learning techniques in cybersecurity," *Energies*, vol. 13, no. 10, p. 2509, 2020.
- [26] X. Liu, J. Zhao, J. Li, B. Cao, and Z. Lv, "Federated neural architecture search for medical data security," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 8, pp. 5628–5636, 2022.
- [27] Z. Zhang, L. Wang, W. Zheng, L. Yin, R. Hu, and B. Yang, "Endoscope image mosaic based on pyramid ORB," *Biomedical Signal Processing and Control*, vol. 71, Article ID 103261, 2022.
- [28] Y. Tang, S. Liu, Y. Deng, Y. Zhang, L. Yin, and W. Zheng, "An improved method for soft tissue modeling," *Biomedical Signal Processing and Control*, vol. 65, Article ID 102367, 2021.
- [29] Y. Liu, J. Tian, R. Hu et al., "Improved feature point pair purification algorithm based on SIFT during endoscope image stitching," *Frontiers in Neurorobotics*, vol. 16, Article ID 840594, 2022.
- [30] S. Liu, B. Yang, Y. Wang, J. Tian, L. Yin, and W. Zheng, "2D/3D multimode medical image registration based on normalized cross-correlation," *Applied Sciences*, vol. 12, no. 6, p. 2828, 2022.
- [31] B. Wixom and H. Watson, "The BI-based organization," *International Journal of Business Intelligence Research*, vol. 1, no. 1, pp. 13–28, 2010.
- [32] B. Wieder and M.-L. Ossimitz, "The impact of Business Intelligence on the quality of decision making—a mediation model," *Procedia Computer Science*, vol. 64, pp. 1163–1171, 2015.
- [33] T. Grublješić and J. Jaklič, "Conceptualization of the business intelligence extended use model," *Journal of Computer Information Systems*, vol. 55, no. 3, pp. 72–82, 2015.
- [34] Z. Lv, D. Chen, and H. Lv, "Smart city construction and management by digital twins and BIM big data in COVID-19

- scenario,” *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 52, 2022.
- [35] S. Ahmad, S. Miskon, T. A. Alkanhal, and I. Tlili, “Modeling of business intelligence systems using the potential determinants and theories with the lens of individual, technological, organizational, and environmental contexts—a systematic literature review,” *Applied Sciences*, vol. 10, no. 9, p. 3208, 2020.
- [36] Y. S. Singh, “Easy designing steps of a local data warehouse for possible analytical data processing,” *ADB Journal of Engineering Technology*, vol. 8, 2019.
- [37] V. Santos, R. Silva, and O. Belo, “Towards a low cost ETL system,” *International Journal of Database Management Systems*, vol. 6, no. 2, pp. 67–79, 2014.
- [38] K. Shaukat, “Student’s performance in the context of data mining,” in *Proceedings of the 2016 19th International Multi-Topic Conference (INMIC)*, IEEE, Islamabad, Pakistan, December 2016.
- [39] M. Z. Latif, “Risk factors identification of malignant mesothelioma: a data mining based approach,” in *Proceedings of the 2020 International Conference on Electrical, Communication, and Computer Engineering (ICECCE)*, IEEE, Istanbul, Turkey, June 2020.
- [40] K. Shaukat, S. Zaheer, and I. Nawaz, “Association rule mining: an application perspective,” *International Journal on Control System and Instrumentation*, vol. 2015, no. 1, pp. 29–38, 2015.
- [41] T. M. Alam, K. Shaukat, I. A. Hameed et al., “An investigation of credit card default prediction in the imbalanced datasets,” *IEEE Access*, vol. 8, pp. 201173–201198, 2020.
- [42] C. Li, M. Dong, J. Li et al., “Efficient medical big data management with keyword-searchable encryption in healthchain,” *IEEE Systems Journal*, vol. 1, 12 pages, 2022.
- [43] T. M. Alam, K. Shaukat, I. A. Hameed et al., “A novel framework for prognostic factors identification of malignant mesothelioma through association rule mining,” *Biomedical Signal Processing and Control*, vol. 68, Article ID 102726, 2021.
- [44] E. Cho, S. A. Myers, and J. Leskovec, “Friendship and mobility: user movement in location-based social networks,” in *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, San Diego, California, USA, August 2011.
- [45] H. Gao, J. Tang, and H. Liu, “Exploring social-historical ties on location-based social networks,” in *Proceedings of the International AAAI Conference on Web and Social Media*, Arizona, USA, November 2012.
- [46] J. Lindqvist, “I’m the mayor of my house: examining why people use foursquare—a social-driven location sharing application,” in *Proceedings of the SIGCHI conference on human factors in computing systems*, Vancouver, BC, Canada, May 2011.
- [47] S. Scellato, “Socio-spatial properties of online location-based social networks,” in *Proceedings of the International AAAI Conference on Web and Social Media*, Catalonia, Spain, July 2011.
- [48] J.-D. Zhang and C.-Y. Chow, “iGSLR: personalized geo-social location recommendation: a kernel density estimation approach,” in *Proceedings of the 21st ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, Orlando, Florida, November 2013.
- [49] G. B. Colombo, “You are where you eat: foursquare checkins as indicators of human mobility and behaviour,” in *Proceedings of the 2012 IEEE International Conference on Pervasive Computing and Communications Workshops*, IEEE, Lugano, Switzerland, May 2012.
- [50] Y. Li, “Exploring venue popularity in foursquare,” in *Proceedings of the 2013 IEEE INFOCOM*, IEEE, Turin, Italy, April 2013.
- [51] S. Lin, “Understanding user activity patterns of the swarm app: a data-driven study,” in *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers*, Maui, Hawaii, September 2017.
- [52] B. Shi, J. Zhao, and P.-J. Chen, “Exploring urban tourism crowding in Shanghai via crowdsourcing geospatial data,” *Current Issues in Tourism*, vol. 20, no. 11, pp. 1186–1209, 2017.
- [53] Z. Gu, Y. Zhang, Y. Chen, and X. Chang, “Analysis of attraction features of tourism destinations in a mega-city based on check-in data mining—a case study of ShenZhen, China,” *ISPRS International Journal of Geo-Information*, vol. 5, no. 11, p. 210, 2016.
- [54] Y. Long, H. Han, Y. Tu, and X. Shu, “Evaluating the effectiveness of urban growth boundaries using human mobility and activity records,” *Cities*, vol. 46, pp. 76–84, 2015.
- [55] C. Wu, X. Ye, F. Ren, and Q. Du, “Check-in behaviour and spatio-temporal vibrancy: an exploratory analysis in Shenzhen, China,” *Cities*, vol. 77, pp. 104–116, 2018.
- [56] J. Wu, J. Li, and Y. Ma, “A comparative study of spatial and temporal preferences for waterfronts in Wuhan based on gender differences in check-in behavior,” *ISPRS International Journal of Geo-Information*, vol. 8, no. 9, p. 413, 2019.
- [57] T. Jia, C. Cai, X. Li, X. Luo, Y. Zhang, and X. Yu, “Dynamical community detection and spatiotemporal analysis in multi-layer spatial interaction networks using trajectory data,” *International Journal of Geographical Information Science*, vol. 36, no. 9, pp. 1719–1740, 2022.
- [58] F. Pezoa, “Foundations of JSON schema,” in *Proceedings of the 25th International Conference on World Wide Web*, Canada, April 2016.
- [59] J. Zhou, Q. Hou, and W. Dong, “Spatial characteristics of population activities in suburban villages based on cellphone signaling analysis,” *Sustainability*, vol. 11, no. 7, p. 2159, 2019.
- [60] M. Lichman and P. Smyth, “Modeling human location data with mixtures of kernel densities,” in *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, New York, NY, USA, August 2014.
- [61] B. P. Loo, S. Yao, and J. Wu, “Spatial point analysis of road crashes in Shanghai: a GIS-based network kernel density method,” in *Proceedings of the 2011 19th international conference on geoinformatics*, IEEE, Shanghai, China, June 2011.
- [62] X. Chen, X. Li, and J. Guan, “Research on the temporal and spatial characteristics of tourist flow in tianchi lake scenic spot based on micro-blog check-in data,” *Areal Research and Development*, vol. 96, 2019.
- [63] Y. Feng, Y. Liu, and X. Tong, “Spatiotemporal variation of landscape patterns and their spatial determinants in Shanghai, China,” *Ecological Indicators*, vol. 87, pp. 22–32, 2018.
- [64] J. Kotus, M. Rzeszewski, and W. Ewertowski, “Tourists in the spatial structures of a big Polish city: development of an uncontrolled patchwork or concentric spheres?” *Tourism Management*, vol. 50, pp. 98–110, 2015.
- [65] C. Lei, A. Zhang, Q. Qi, H. Su, and J. Wang, “Spatial-temporal analysis of human dynamics on urban land use patterns using social media data by gender,” *ISPRS International Journal of Geo-Information*, vol. 7, no. 9, p. 358, 2018.
- [66] Q. Liu, L. Hou, S. Shaukat, U. Tariq, R. Riaz, and S. S. Rizvi, “Perceptions of spatial patterns of visitors in urban green spaces for the sustainability of smart city,” *International*

- Journal of Distributed Sensor Networks*, vol. 17, no. 8, Article ID 155014772110340, 2021.
- [67] S. Scellato, "Socio-spatial properties of online location-based social networks," in *Proceedings of the Fifth international AAAI conference on weblogs and social media*, Catalonia, Spain, July 2011.
- [68] B. Wang, F. Zhen, Z. Wei, S. Guo, and T. Chen, "A theoretical framework and methodology for urban activity spatial structure in e-society: empirical evidence for Nanjing City, China," *Chinese Geographical Science*, vol. 25, no. 6, pp. 672–683, 2015.
- [69] H. Ullah, Wan, Haidery, Khan, Ebrahimpour, and Luo, "Analyzing the spatiotemporal patterns in green spaces for urban studies using location-based social media data," *ISPRS International Journal of Geo-Information*, vol. 8, no. 11, p. 506, 2019.
- [70] H. Ullah, W. Wan, S. A. Haidery, N. U. Khan, Z. Ebrahimpour, and A. A. M. Muzahid, "Spatiotemporal patterns of visitors in urban green parks by mining social media big data based upon WHO reports," *IEEE Access*, vol. 8, pp. 39197–39211, 2020.
- [71] Z. Ebrahimpour, Wan, Cervantes, Luo, and Ullah, "Comparison of main approaches for extracting behavior features from crowd flow analysis," *ISPRS International Journal of Geo-Information*, vol. 8, no. 10, p. 440, 2019.
- [72] Q. Liu, H. Ullah, W. Wan et al., "Analysis of green spaces by utilizing big data to support smart cities and environment: a case study about the city center of shanghai," *ISPRS International Journal of Geo-Information*, vol. 9, no. 6, p. 360, 2020.
- [73] S. Ali Haidery, H. Ullah, N. U. Khan, K. Fatima, S. S. Rizvi, and S. J. Kwon, "Role of big data in the development of smart city by analyzing the density of residents in shanghai," *Electronics*, vol. 9, no. 5, p. 837, 2020.
- [74] Q. Liu, H. Ullah, W. Wan et al., "Categorization of green spaces for a sustainable environment and smart city architecture by utilizing big data," *Electronics*, vol. 9, no. 6, p. 1028, 2020.
- [75] L. Hou, Q. Liu, J. Nebhen, M. Uddin, M. Ullah, and N. U. Khan, "Analyzing the check-in behavior of visitors through machine learning model by mining social network's big data," *Computational and Mathematical Methods in Medicine*, vol. 2021, 11 pages, 2021.
- [76] V. Heikinheimo, H. Tenkanen, C. Bergroth, O. Jarv, T. Hiippala, and T. Toivonen, "Understanding the use of urban green spaces from user-generated geographic information," *Landscape and Urban Planning*, vol. 201, Article ID 103845, 2020.
- [77] G. Boeing, "Spatial information and the legibility of urban form: big data in urban morphology," *International Journal of Information Management*, vol. 56, Article ID 102013, 2021.
- [78] W. Wu, "City profile: shanghai," *Cities*, vol. 16, no. 3, pp. 207–216, 1999.
- [79] T. Nam and T. A. Pardo, "Conceptualizing smart city with dimensions of technology, people, and institutions," in *Proceedings of the 12th annual international digital government research conference: digital government innovation in challenging times*, College Park, Maryland, USA, June 2011.
- [80] W. Zheng, Y. Xun, X. Wu, Z. Deng, X. Chen, and Y. Sui, "A comparative study of class rebalancing methods for security bug report classification," *IEEE Transactions on Reliability*, vol. 70, no. 4, pp. 1658–1670, 2021.
- [81] W. Yang, X. Chen, Z. Xiong, Z. Xu, G. Liu, and X. Zhang, "A privacy-preserving aggregation scheme based on negative survey for vehicle fuel consumption data," *Information Sciences*, vol. 570, pp. 526–544, 2021.
- [82] N. U. Khan, W. Wan, S. Yu, A. A. M. Muzahid, S. Khan, and L. Hou, "A study of user activity patterns and the effect of venue types on city dynamics using location-based social network data," *ISPRS International Journal of Geo-Information*, vol. 9, no. 12, 733 pages, 2020.
- [83] N. U. Khan, W. Wan, and S. Yu, "Spatiotemporal analysis of tourists and residents in Shanghai based on location-based social network's data from Weibo," *ISPRS International Journal of Geo-Information*, vol. 9, no. 2, 70 pages, 2020.

## Research Article

# Trends in Intelligent and AI-Based Software Engineering Processes: A Deep Learning-Based Software Process Model Recommendation Method

Fahad H. Alshammari 

*College of Computing and Information Technology, Shaqra University, Shaqra, Saudi Arabia*

Correspondence should be addressed to Fahad H. Alshammari; [fahad.h@su.edu.sa](mailto:fahad.h@su.edu.sa)

Received 24 August 2022; Revised 17 September 2022; Accepted 20 September 2022; Published 5 October 2022

Academic Editor: Ateeq Ur Rehman

Copyright © 2022 Fahad H. Alshammari. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In recent years, numerous studies have successfully implemented machine learning strategies in a wide range of application areas. Therefore, several different deep learning models exist, each one tailored to a certain software task. Using deep learning models provides numerous advantages for the software development industry. Testing and maintaining software is a critical concern today. Software engineers have many responsibilities while developing a software system, including coding, testing, and delivering the software to users *via* the cloud. From this list, it is easy to see that each task calls for extensive organization and preparation, as well as access to a variety of resources. A developer may consult other code repositories, websites with related programming content, and even colleagues for information before attempting to build and test a solution to the problem at hand. In this investigation, we aim to identify the factors that led to developing the recommender. This system analyzes the recommender's performance and provides suggestions for improving the software based on users' opinions.

## 1. Introduction

When developing a software system, software engineers execute various tasks, including creating code, testing code, deploying to the cloud, and coordinating *via* e-mail and meetings [1]. Each of these tasks necessitates searching for and working with a wide range of information and resources, as well as planning and preparing for the upcoming one [2]. A developer may investigate other code repositories for prospective solutions, explore online sites with relevant programming material, or contact coworkers for information before programming a possible solution to the problem at hand and testing the answer [3].

For example, performing these tasks can be intimidating for novices in the field [4]. Near-perfect performance in these activities is nearly unattainable for even the most experienced coders. Recommender systems for software engineering have been implemented to easily perform tasks and improve workflow [5]. In other words, “software

applications that deliver information items deemed to be relevant for software engineering tasks” are “recommenders” for the discipline [6]. Software engineers are used to working with certain recommenders that are closely relevant to their development operations. Such issues as missing import declarations in Java code can be solved using recommenders in various integrated development environments, such as the Eclipse IDE4 [7].

Recommendation systems for different tasks and workflows have been developed, including those for code reorganization, learning the next set of commands, and discovering needs. For instance, the Eclipse Mylyn recommender, which provides specific recommendations of which source code is connected with a task, has been demonstrated to boost the productivity of developers. Recommenders have much unrealized potential in the software development process because of their vast variety of actions [8].

One of the primary problems with the current recommender system is that it forecasts products that the user will



find irrelevant or uninteresting. As a result, a recommender system is required, which must supply services in accordance with the resemblance of goods. By incorporating user and product data into a collaborative recommendation system, true user preferences can be learned [9–11].

The first stage in developing a recommender is to define the problem the recommender is intended to solve and verify the assumption that a recommender can deliver suggestions of value to the developer facing the problem. Framing the problem is the term we use to describe the activities occurring during this phase. The introduction's definition of a software engineering recommender provides a foundation for investigating the problem and solution targeted by a recommendation engine. The task and context for which a recommender will be used must be crystal apparent when thinking about creating one. Another consideration is for whom a recommender is intended: developers or end users. The idea of a task targeted by a recommender relates to the specific purpose of a developer at a certain moment in time, such as the implementation of an assigned feature in source code. Even though a developer is aware of the current task at all times, the task may not be expressed directly in the code. The context of a recommender refers to the information and tool environment in which the task is conducted, such as the source code and other artifacts available and the set of tools that can be used to complete the work. The context also captures the developer's steps in completing the given task. This helps define when and what information a recommender may provide: novices often have fundamentally different information needs compared to experts. While frequent proposals may be helpful to the first group, the latter often has a poor tolerance for interruptions of their work that convey already known facts. The main contributions of this study are as follows:

- (a) Determining the inputs for the recommender's construction was how we phrased the issue
- (b) This system provides recommendations for software development based on client happiness and evaluates the usefulness of the recommender

## 2. Related Work

Wen et al. [12] systematically examined machine learning models from four perspectives: the kind of ML approach, estimation accuracy, model comparison, and context of estimation, which is the goal of this study. A systematic review of empirical studies on the ML model published between 1991 and 2010 was conducted. The author compiled a list of 84 primary studies related to our research question and detected eight different types of ML approaches used in SDEE models after looking into these studies. Overall, these ML models have better estimation accuracy than non-ML models and are near to it. For this reason, certain ML models are more effective in certain estimation scenarios. SDEE is a potential field for ML models. However, the industry's use of ML models is still limited, necessitating additional efforts and financial incentives. Following the conclusions of this review, the author offers advice for researchers and guidance for practitioners.

Wan et al. [13] were curious about the impact of machine learning on software development techniques, given the growing popularity of this approach. From interviews with 14 people and surveys with 342 people from 26 nations across four continents, we could identify substantial differences between the development of machine and non-machine learning systems. Software engineering (e.g., requirements, design, testing, and process) and work characteristics are significantly different across the two groups, according to our research (e.g., skill variety, problem-solving, and task identity). In light of our findings, the author outlined potential future research areas and offered practice-oriented suggestions.

Del Carpio and Angarita [14] used machine learning approaches in various knowledge domains with promising results. Many deep learning models now focus on a wide range of software operations, which is a good sign for the future systematic investigation of deep learning model-supported software processes that yield useful findings for the software industry. Software testing and maintenance were the most often studied subprocesses in this study. It is common to utilize deep learning models such as CNN and RNN to process bug reports, malware categorization, and recommendation creation in these subprocesses. Some solutions are focused on estimating effort, classifying software requirements, identifying GUI visual aspects, identifying code authors, finding the similarity between source codes, predicting and classifying defects, and analyzing bug reports in testing and maintenance operations.

Meziane and Vadera [15] suggested that, due to its ability to automate time-consuming or complex processes, artificial intelligence has recently gained much attention. There have been no exceptions to this rule regarding software engineering projects. Artificial intelligence and software maintenance are covered in depth in this thesis. The recent advances in applying artificial intelligence to software maintenance duties were also studied through thorough mapping research. Research kind, research contribution, software maintenance domains, and artificial intelligence solution type were the most important aspects of this study.

Barenkamp et al. [16] involved a systematic evaluation of prior research and five qualitative interviews with software developers. The study's conclusions are categorized throughout software development. Major AI achievements and future potentials include (a) using algorithms to automate time-consuming, routine tasks in software development and testing (such as bug hunting and documentation); (b) conducting structured analyses of large datasets to uncover patterns and new information clusters; and (c) conducting systematic evaluations of these datasets in neural networks. AI accelerates development, reduces expenses, and increases efficiency. Software engineering automation is superior to the present AI, which relies on human-made structures and is essentially reproductive. Developers can enhance their creativity with AI tools.

Harman claimed that the artificial intelligence (AI) approaches to software engineering also focuses on the software development related challenges [17]. While search-based software engineering is a more recent development,

the field's history of work in probabilistic reasoning and machine learning for software engineering is well-established. For the purpose of this paper, the author examined some of the connections between these two areas of research, claiming that they share many characteristics.

Tate [18] compared software quality models. Case studies apply software quality models to the current processes. Case study results complement empirical model assessment. Standard selection criteria are used to recommend and select models. Procedures are evaluated using success criteria. Theoretical assessment methods evaluate process model quality. Conformity to ideal process quality model requirements and relevance to software stakeholders are tested. Discussing the models' breadth and scale: empirical assessment methods are established to evaluate the model's performance in real software operations. There are approaches to determine if process quality models produce different results and, if so, which model to choose. Case study software processes are measured for differences.

Fadhil et al. [11] determined how AI can improve software issue detection and prediction methods. Artificial intelligence has helped identify software issues and predict bugs, as data shows. Combining AI with software engineering reduces overhead and produces more efficient solutions, improving software quality.

Kothawar and Vajrapu [19] addressed these behaviors' difficulties and solutions. Methods: the author chose 15 best practices from eight startups, each with unique challenges and solutions. Our research indicates startups' mixed prioritization. Six of the eight companies used formal methods, while two used unstructured prioritization. Startups' value: prioritizing based on consumer input and ROI is key. This study examines startup priority needs and obstacles. The literature supports the study's findings. Finding solutions helps practitioners. The poll should include Swedish software startups. Some of these solutions may also be useful for practitioners wishing to begin a software startup and priority requirements.

This study's aggregation method is clear, realistic, and interpretable [9]. This method makes quality model and metric-based software quality assessment reliable and reproducible. Based on all observable software artifacts, good and bad quality are assigned probabilities. Validation was theoretical and empirical. Bug prediction, maintainability, and information quality were evaluated. Software visualization was used to evaluate the usefulness of aggregation for multivariate data and the impact of different aggregation methods. Finally, the author assessed MCR's transferability and used it to rate real-world options. The author used machine learning, created a benchmark employing regression issues, and evaluated how well the aggregate result matches a ground truth and represents input variables. Our method is accurate, sensitive, and facilitates multicriteria decision-making. Our approach can be used as an agnostic unsupervised predictor without ground truth.

Recently, sentiment analysis on social networks, such as Twitter and Facebook, has become a valuable tool for gaining insight into the thoughts and feelings of people. In

contrast, sentiment analysis suffers from the difficulties of natural language processing (NLP). Deep learning models have recently been a promising solution to NLP difficulties. To address the issues with sentiment analysis, such as sentiment polarity, the paper [10] analyzes the most recent experiments to make use of deep learning. Word embedding and the TF-IDF model have been used to analyze several different datasets. Comparative studies of the experimental findings for various models and input features have also been undertaken.

Software defect prediction anticipates troublesome code sections to help find faults and priorities testing. Previous work focused on manually encoding program information and using machine learning to generate accurate prediction models. Standard characteristics do not capture semantic differences between programs for accurate prediction models [8]. Deep learning is proposed to bridge the gap between program semantics and fault prediction characteristics. The deep belief network (DBN) learns semantic features from Abstract Syntax Tree (AST) token vectors automatically. Our research on 10 open-source projects shows that our automatically learned semantic features increase both within-project and cross-project defect prediction over traditional characteristics. Precision, recall, and F1 improve WPDP by 14.7%, 11.5%, and 14.2%, respectively. Our semantic feature-based technique beats TCA + by 8.9% in F1 for CPDP.

Reference [20] proposed LEMNA, a high-fidelity security explanation approach. LEMNA generates a limited set of features that explain how an input sample is categorized. The goal is to create a simple interpretable model to approximate the deep learning decision boundary. It manages feature dependency to better interact with security applications (such as binary code analysis) and nonlinear local boundaries to boost explanation fidelity. Local interpretable model (LIM): the author tested our method with two deep learning security apps (a malware classifier and a function start detector for binary reverse engineering). Extensive testing demonstrates that LEMNA's explanation is more correct than others. The author shows how LEMNA may help machine learning developers verify model behavior, fix classification issues, and automatically patch target model defects.

Reference [7] reviewed machine learning papers for software project management. Web Science, Science Directs, and IEEE Explore have research on machine learning, software project management, and methodology. Three repositories contain 111 papers in four groupings. First group: software project management papers. The second category contains machine learning methods and tactics utilized in projects. The third category comprises studies on machine learning management phases and tests, as well as study findings, contribution to and promotion of machine learning project prediction, and other studies. It gives a broader context for future project risk management efforts. Machine learning-based project risk assessment is more successful in reducing project losses, increasing project success, and reducing project failure probabilities while increasing the growth output ratio.

Recent machine learning discoveries have prompted interest in integrating AI into IT software and services. To fulfill this goal, organizations adapted their development methodologies. The author shares research on Microsoft's AI-app development teams. It is built on designing AI apps (search and NLP) using data science tools (R and Python) (e.g., application diagnostics and bug reporting). Reference [5] found that multiple Microsoft teams have integrated this workflow into established, well-evolved software engineering processes, providing insights into numerous important engineering problems organizations may encounter while developing large-scale AI products for the market. These difficulties required Microsoft's best practices. Aside from that, the author found three main AI differences: (1) model customization and reuse demand different abilities than those found in software teams. (2) AI components are more challenging to handle as independent modules than typical software components. Microsoft teams provided critical knowledge.

Yang et al. [6] proposed "deep neural networks" (DNNs) and an updated model training approach. Alpha Go showed deep learning's potential in 2016. Deep learning helps software engineering (SE) experts construct cutting-edge research tools. Model selection, internal structure, and tuning affect DNN performance in SE. Deep learning in SE is understudied. The author searched for relevant publications since 2006. First, SE deep learning is shown. SE's deep learning methods are classified. The author looked at deep learning model optimization methodologies and highlighted SE research problems that will benefit from DNNs. Our findings highlight existing problems and suggest a potential study route.

Machine learning is rapidly used by the software engineering community as a means of transforming modern software into intelligent and self-learning systems. Software engineers are still exploring methods in which machine learning can aid with various stages of the software development life cycle. Herein, the author reports the results of a study on the application of machine learning at various stages of the software development life cycle. Overall, [3] investigated the relationship between software development life cycle stages and machine learning tools, techniques, or types, which is a broad goal. In an attempt to answer the question of whether machine learning favors specific stages or methodologies, we conduct a comprehensive analysis.

Business transactions, revenues, and general success are becoming increasingly dependent on the use of recommendation systems. Recommendation systems and their implementation approaches are the focus of this survey. The components and attributes of a recommender system can change based on the organization's needs. Design criteria and key recommender system attributes are presented in this study. There are a few well-known approaches that are scrutinized. In conclusion, [4] introduced movie recommenders from the three most relevant industries: film, music, and online shopping. The survey seeks to provide readers with a broad understanding of the circumstances in which certain recommender systems are appropriate.

Machine learning models are frequently developed by data scientists to handle a wide range of problems in both industry and academia, but they are not without their own set of hurdles. One of the issues with machine learning development is that many people working in the field are unaware of the benefits that may be reaped from following the steps outlined in the software engineering development lifecycle (SEDL). Of course, because machine learning systems are distinct from typical software systems, there will be certain peculiarities in the development process. Regarding software engineering, [2] aimed to examine the issues and practices that arise during model creation by looking at how developers might benefit from using or changing the standard workflow to machine learning.

Software engineering has recently used deep learning (SE). Unanswered questions remain. Li et al. [1] looked at 98 SE publications that employ deep learning to tackle these questions. Deep learning technologies have simplified 41 SE jobs across all phases. Deep learning models and their variations are utilized to answer 84.7% of SE issues in publications. Deep learning's practicality is questioned. More SE scholars may be interested in improving deep learning-based solutions in the future.

### 3. Methodology

In this section, we have proposed a novel framework of LSTM which can recommend the software development features based on the dataset of clients. Figure 1 shows the proposed framework workflow of the current study:

*3.1. Dataset Description.* The dataset used in this study is an excel-generated synthetic dataset curated from a real BI tools' dataset. This dataset has 100 rows and 11 features with 1 output feature (i.e., rating); when the rating of software is more than 3, this will be recommended. Otherwise, it will not be recommended by the proposed model. Table 1 shows the dataset description and feature explanation.

Table 2 shows the dataset samples from the acquired dataset as given below.

Figure 2 shows the visualization of the dataset and frequency distribution of each feature as given below.

Figure 3 shows the distribution of feature business scale with respect to large, small, and medium deployment on premise, hybrid, and cloud OS for Windows, Mac, and Linux and pricing on Freemium, open source, and enterprise.

*3.2. Raw Data Processing.* The raw data have been collected. Finally, data purification has been completed using various methods, such as deleting duplicates and null values. This technique is employed in data mining to transform unstructured data into a form suitable for analysis. It is not uncommon for data in the real world to be inconsistent or even missing. Prediction models are complicated when classifications are not dispersed uniformly throughout. The number of occurrences in each class is often the same in categorization machine learning algorithms. In the wake of this study, resampling procedures have substantially evolved.

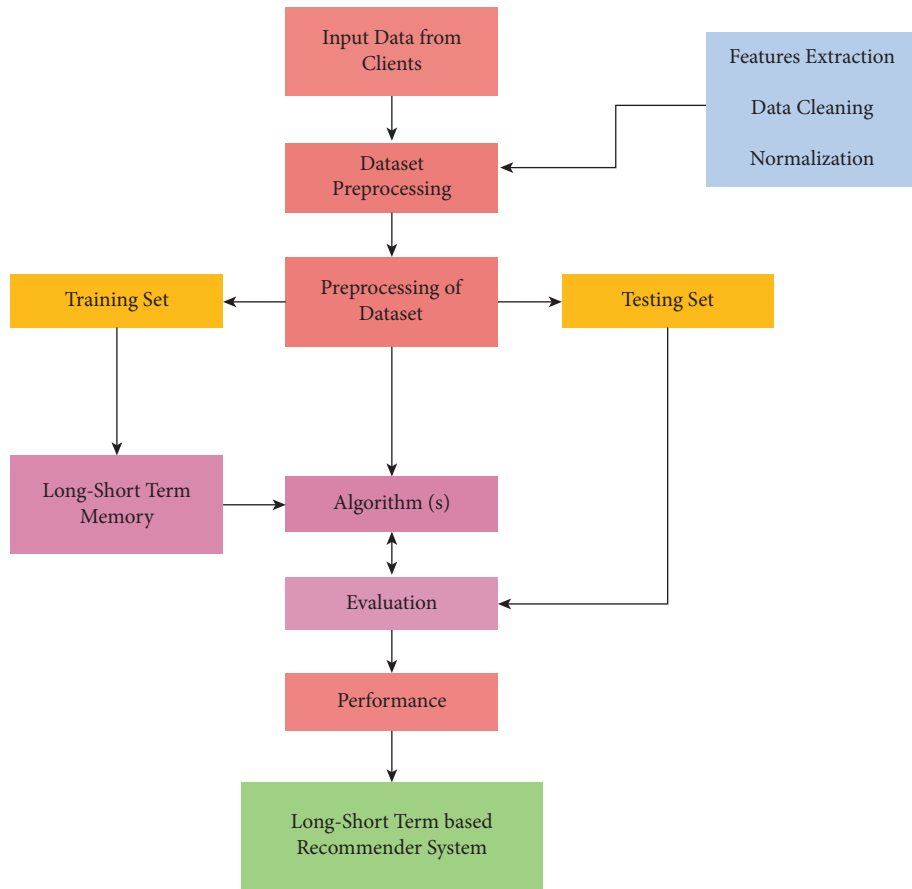


FIGURE 1: The proposed framework workflow.

TABLE 1: Dataset description and feature explanation.

Features	Description	Variables type
Category	Category comprises the type of BI tool, as well as the industry in which it can be used	Input variable
Business scale	This identifies the size of the company that the BI tool is designed to serve, such as small, medium, or large	Input variable
User type	Whether they are a business user or an analyst with data science skills, this indicates the sort of user (not all BI tools are easy to use and not all tools possess powerful data processing capabilities)	Input variable
No of users	This offers information on how the BI tool may be implemented, such as cloud, on-premise, or hybrid OS: this specifies the type of operating system necessary for the tool installation	Input variable
Pricing	Pricing: this reveals whether the program has a freemium version or an enterprise edition for data visualization on mobile devices	Input variable
Ratings	On a scale of 5.0, users rate this product	Output variable

TABLE 2: Dataset samples from the acquired dataset.

Category	Industry	Business scale	User type	No. of users	Deployment	OS	Mobile apps	Pricing	Rating
100001	Data management	Utilities	Large	Business	Single	Cloud	Linux	Y	4.5
100002	Database/ERP	Food	Large	Business	Single	Premise	Mac	Y	5.0
100003	Data analysis	Manufacturing	Large	Business	Single	Premise	Linux	N	5.0
100004	Data analysis	IT	Medium	Business	Multiple	Premise	Mac	Y	4.3
100005	Benchmarking	Food	Medium	Analyst	Multiple	Cloud	WIN	N	4.7

Remove records from each cluster such that the majority class records are captured and undersampling is prevented. For more diverse synthetic samples, oversampling can be utilized in place of producing identical reproductions of data from the

minority classes [21]. When conducting data mining research, it is critical that our dataset is balanced and consistent. It is possible to find outliers in a dataset. An outlier in a dataset is a value that stands out from the rest because of its uniqueness.

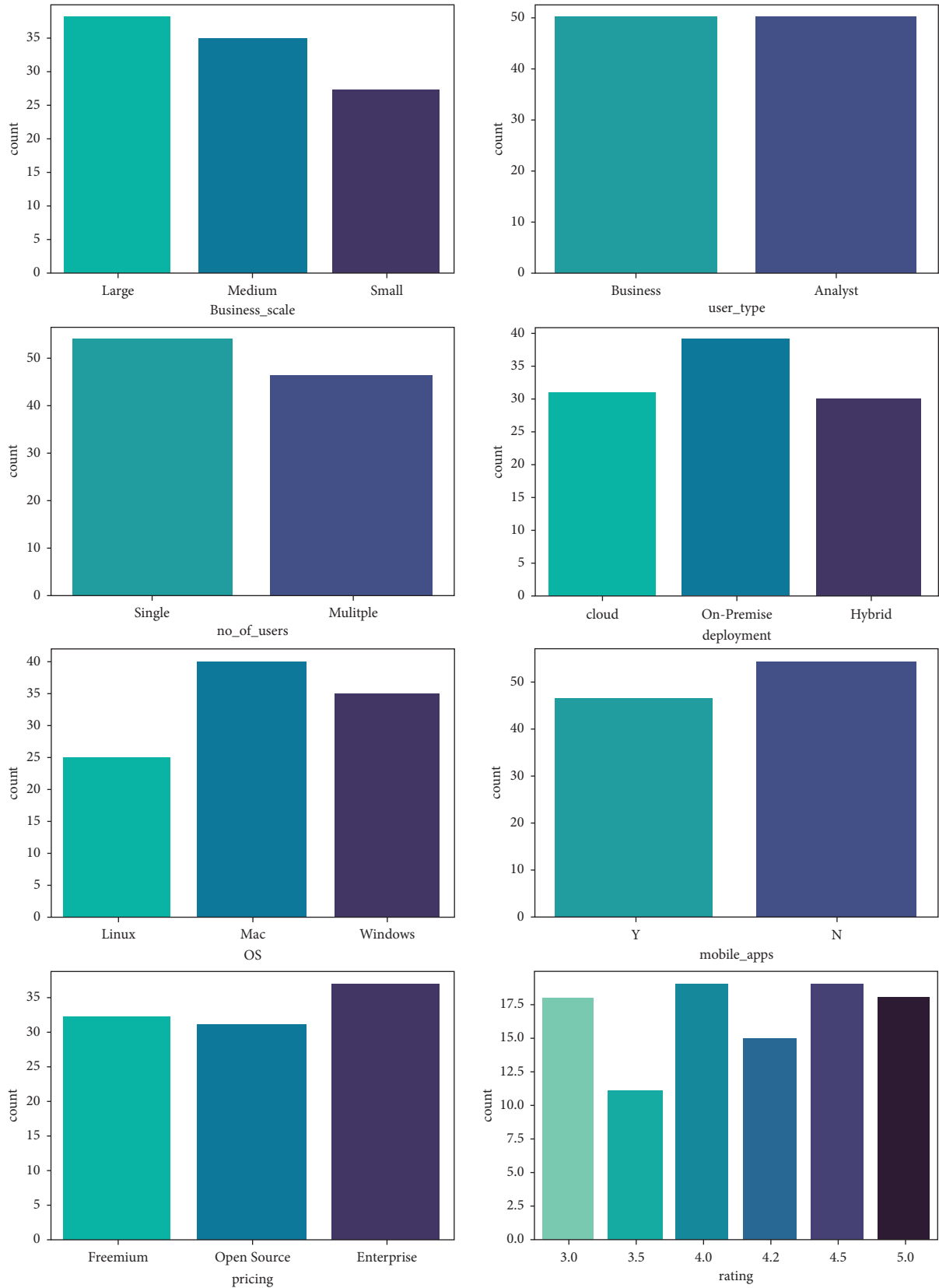


FIGURE 2: Visualization of the dataset and frequency distribution of attributes.

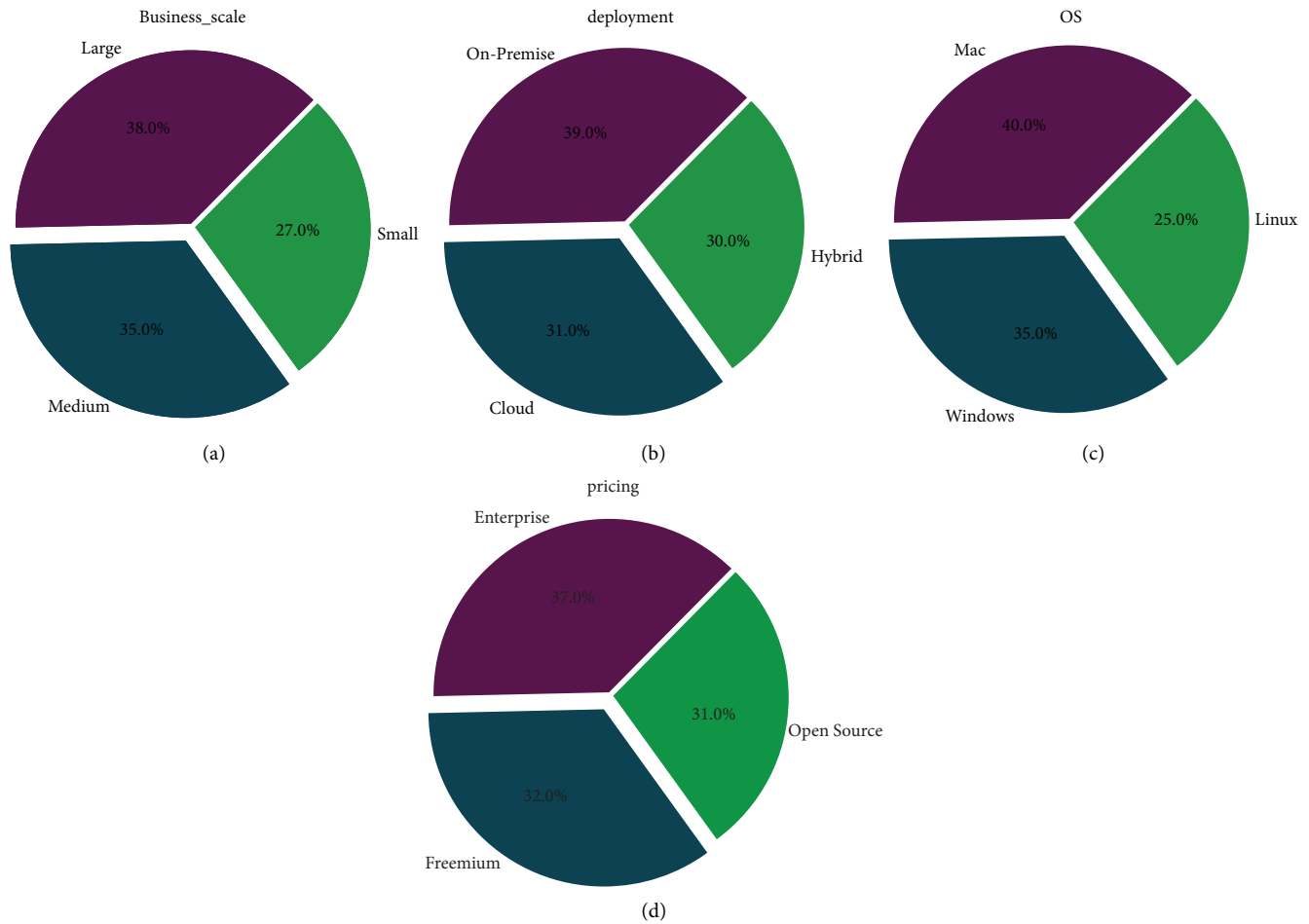


FIGURE 3: Distribution of features: (a) business scale, (b) deployment, (c) OS, and (d) pricing.

The outliers could result from reading errors, equipment faults, or human error. Before undertaking any statistical analysis or study, it must be deleted from the dataset. The analysis and subsequent treatment can be influenced by incomplete or erroneous findings from any information outlier [22, 23].

**3.3. Feature Engineering.** By using data from a certain domain, learning machines can use these functions. In order to make machine learning representations of raw data, this must be done manually. Correlation matrices are used in this study to determine the correlation between the variables. Covariance matrices are the same as correlation matrices. Using the correlation, one may determine the strength of a linear link. The concept of correlation summarizes the frequency and direction of a straight-line link between two quantitative variables. Values can be represented by  $r$ , which ranges from  $-1$  to  $+1$ .

**3.4. Proposed Model.** In the proposed model (shown in Figure 4), input sequences are feature embedded and then extracted in the contented layer. There is a hyperband optimization algorithm that can be used to distribute hyperparameter tuning for TensorFlow models in just a few lines

of code in the Keras-Tuner module. For hyperparameter tuning, a validation dataset containing 10% of randomly selected samples from the training data is used. Furthermore, we employed sparse categorical accuracy as a ranking metric for optimization trials. We experimented with various batch size variables before settling on batch size = 512. Data from previous optimization stages are used to train a final model with a set of hyperparameters that is as good as it can possibly be. In order to assess the accuracy of our new recommender system, we implemented a back-testing technique.

**3.4.1. Novel LSTM Cell.** Long short-term memory networks are a subset of the broader category of recurrent neural networks. An example of time- or sequence-dependent behavior is language, stock prices, and power demand; recurrent neural networks seek to represent such phenomena. In order to achieve this, the output of a layer in a neural network at time  $t$  is fed back into the input of the same layer at time  $t + 1$ . Figure 5 shows the modified recurrent units of the new version of LSTM:

During training and prediction, recurrent neural networks are “unrolled” programmatically, resulting in Figure 6.

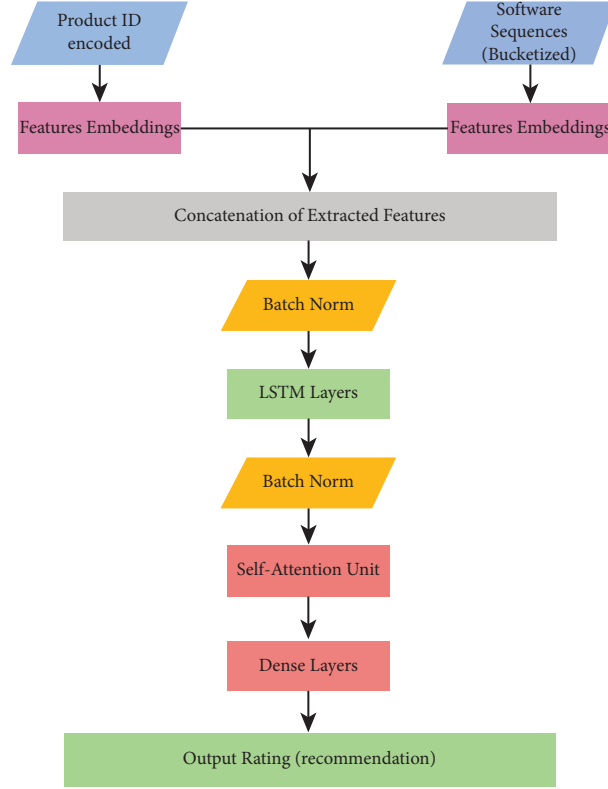


FIGURE 4: Proposed model architecture.

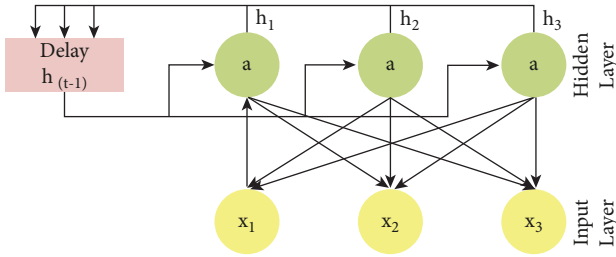


FIGURE 5: Recurrent nodes of modified LSTM.

New data are sent to the network at each time step, and the output of the previous  $F$  ( $\$h_{t-1}$ ) is also supplied, as shown in Figure 6.

In place of the typical neural network layers, an LSTM network uses LSTM cell blocks to store information for future use. The input, forget, and output gates are all parts of these cells that will be discussed in greater depth below. Our planned LSTM cell is depicted graphically below in Figure 7.

**3.4.2. Input Gate.** First, a  $\tanh$  activation function is applied, compressing the input to a range from  $-1$  to  $1$ . To put it another way,

$$g = \tanh(b^g + x_t^{U^g} + h_{t-1}V^g), \quad (1)$$

where  $x_t^{U^g}$  and  $V^g$  represent the input and previous cell output weights and  $b^g$  represents the input bias. The  $g$  exponents do not represent an increased power but rather

the weights and biases used in the input calculations (as opposed to the input gate, forget gate, output gate etc.). The output of the input gate, which is a chain of sigmoid-activated nodes, is multiplied by this compressed input, element by element:

$$i = a\left(\left(b^i + x_t^{U^i} + h_{t-1}V^i\right)\right). \quad (2)$$

**3.4.3. Forget Gate and State Loop.** Forget gate of the cell is expressed as

$$f = a\left(\left(b^f + x_t^{U^f} + h_{t-1}V^f\right)\right). \quad (3)$$

The product of the previous state with the forget gate yields an expression of the form  $((b^f + x_t^{U^f} + h_{t-1}V^f))$  as its output. Following the forget gate/state loop, the product is

$$s_t = s_{t-1} \times f \times g. \quad (4)$$

**3.4.4. Output Gate.** The output gate of LSTM is expressed as

$$O = a\left(\left(b^o + x_t^{U^o} + h_{t-1}V^o\right)\right). \quad (5)$$

Finally, the product of all gates is

$$h_i = \tanh\left(a\left(\left(b^{o+f+s+i} + x_t^{U^{o+f+s+i}} + h_{t-1}V^{o+f+s+i}\right)\right)\right). \quad (6)$$

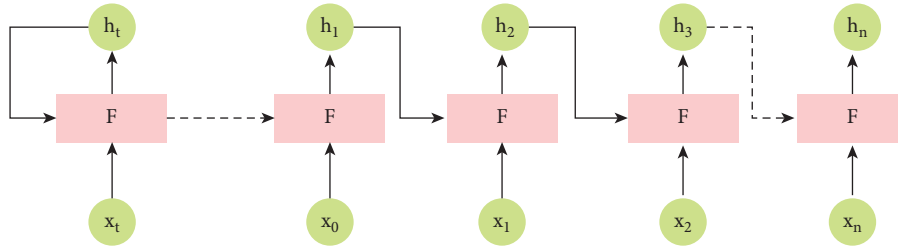


FIGURE 6: Unrolled nodes.

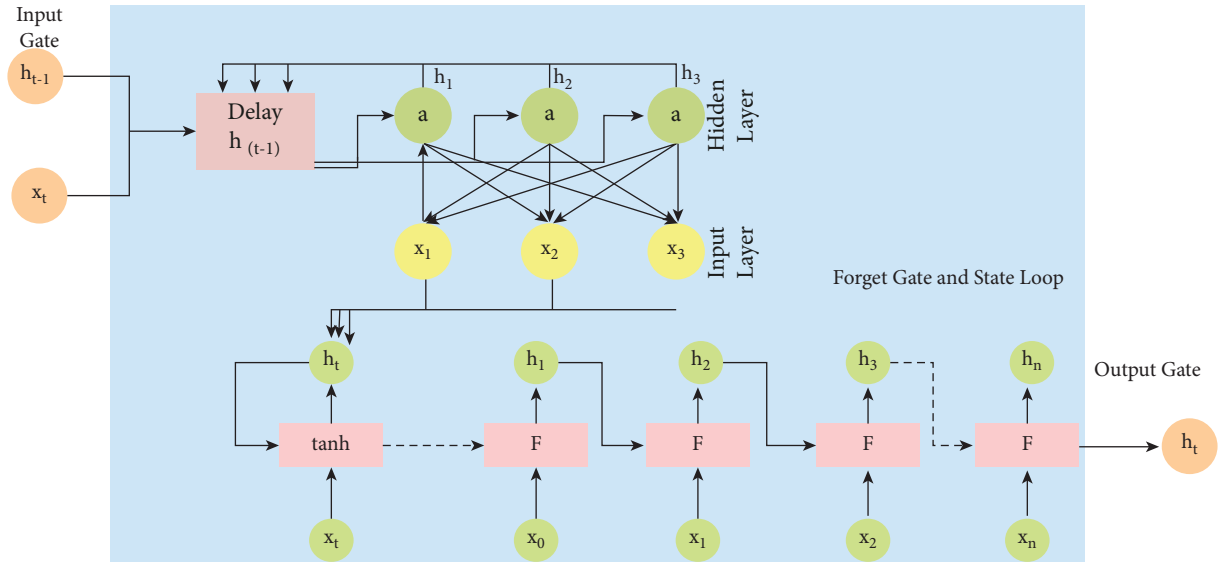


FIGURE 7: Modified cell of LSTM.

TABLE 3: Description of metrics.

Metric	Description
Accuracy	$Accuracy = TP / (TP + TN) * 100$
Recall	$Recall = \sum recommender / \sum (new - known) * 100$

Recall and accuracy were utilized to assess the effectiveness of the strategies under consideration for the software development recommender system. The computations of the metrics utilized in this study are shown in Table 3.

#### 4. Results and Discussion

Our approach was put to the test using data from the Steam project. In order to test our strategy, there are no existing datasets that can be used for this purpose. For testing purposes, we used the most recent records as a test set and the rest of the records as training sets.

In this experiment, we used a serial filling with a time series length of  $T=12$  and a dimensionality reduction with an aimed dimension  $k=50$ . Finally, we gave each user a list of the top 50 ( $N=50$ ) things. We used two separate control trials to assess the effectiveness of each component of our strategy. Neither the serial filling (noSF) nor the dimensionality reduction (noDR) was applied in one experiment.

TABLE 4: Recall rates of the proposed LSTM with different approaches.

Approach	Recall rate
IR	0.0834
IF	0.054
TD	0.012

TABLE 5: Time of the proposed LSTM with different approaches.

Approach	Training time (seconds)	Testing time (seconds)
IR	1400	15
IF	1150	12

In order to establish a baseline, we compared our method to collaborative filtering for implicit feedback (IF) and temporal decay (TD). To evaluate the correctness of our recommendations, we looked at the recall rate, whereas for determining system efficiency, we looked at training and execution times. Our final step was to examine each software's average recommendation time to see if there was a wide range of recommendation times for each method.

The recall rates for various techniques are shown in Table 4. When IR and serial filling were tested, it was found



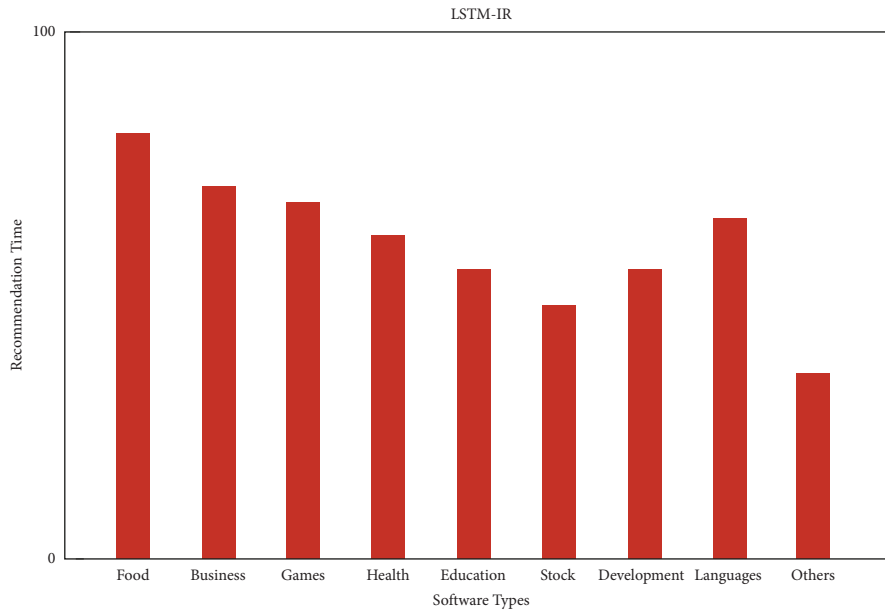


FIGURE 8: Recommendation of LSTM with IR.

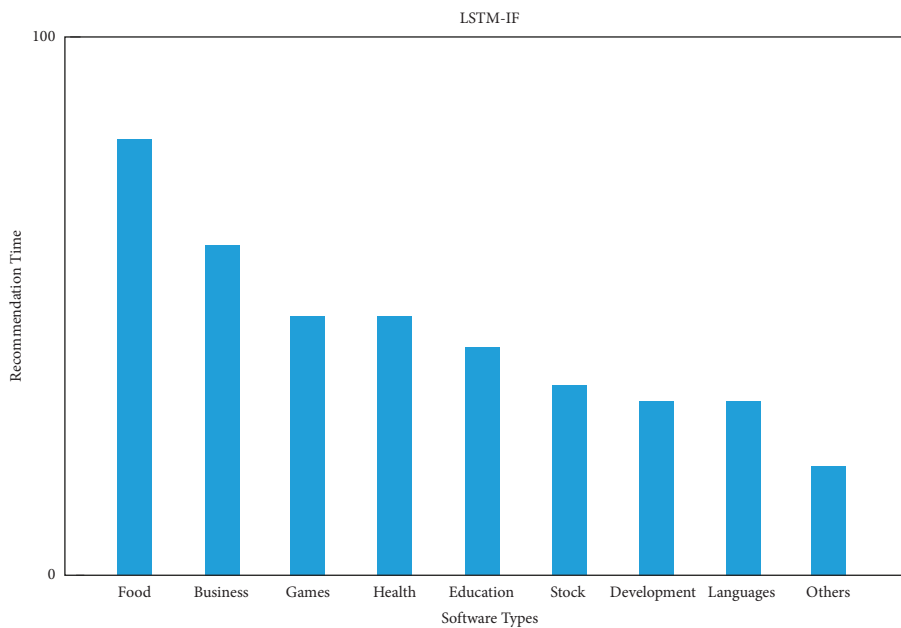


FIGURE 9: Recommendation of LSTM with IF.

to have a greater recall rate than the baseline techniques. Time spent in IR is shown in Table 5. Matrix factorization is a useful way to reduce the number of dimensions in a system because the recall rate of IR was nearly the same as that of IRnoDR.

Figure 8 displays the IR and Figure 9 depicts the IF distribution of the top software recommendation times. Because IR recommends more diverse items than baseline collaborative filtering, we can conclude that our approach is more diverse than baseline collaborative filtering.

## 5. Conclusions

Within the scope of this work, an LSTM-based recommendation model for interaction records was suggested. Based on the results of our evaluations, our model performed admirably in all three categories: accuracy, efficiency, and variety. In the future, we intend to evaluate the generalizability of our approach by applying it to a wide variety of datasets. In addition, we considered the total amount of time spent communicating with one another as a quality factor in this study. There is a high probability that reviews

will be distorted due to the viewpoints of various individuals and types of goods. As a direct consequence of this, we ought to direct our attention going forward toward enhancing the quality of our rating vectors in the future. In order to deal with time series, we will also investigate a variety of other approaches and models.

## Data Availability

The data used to support the findings of this study are included within this article.

## Conflicts of Interest

The author declares no conflicts of interest.

## Acknowledgments

This study could not have been started or achieved without the encouragement of Shaqra University and its continued support. The research work was supported by the College of Computing and Information Technology, Shaqra University, KSA.

## References

- [1] X. Li, H. Jiang, Z. Ren, G. Li, and J. Zhang, "Deep learning in software engineering," 2018, <https://arxiv.org/ftp/arxiv/papers/1805/1805.04825.pdf>.
- [2] G. Lorenzoni, P. Alencar, N. Nascimento, and D. Cowan, "Machine learning model development from a software engineering perspective: a systematic literature review," 2021, <https://arxiv.org/abs/2102.07574>.
- [3] S. Shafiq, A. Mashkoo, C. Mayr-Dorn, and A. Egyed, "A literature review of using machine learning in software development life cycle stages," *IEEE Access*, vol. 9, pp. 140896–140920, 2021.
- [4] N. Koneru, S. Rai, S. S. kumar, and S. Koppu, "Deep learning-based automated recommendation systems: a systematic review and trends," *Turkish Journal of Computer Mathematics Education*, vol. 12, no. 6, pp. 3326–3345, 2021.
- [5] S. Amershi, A. Begel, C. Bird et al., "Software engineering for machine learning: a case study," in *Proceedings of the 2019 IEEE/ACM 41st International Conference on Software Engineering: Software Engineering in Practice (ICSE-SEIP)*, pp. 291–300, Montreal, QC, Canada, May 2019.
- [6] Y. Yang, X. Xia, D. Lo, and J. Grundy, "A survey on deep learning for software engineering," *ACM Computing Surveys*, vol. 54, no. 10, 2022.
- [7] M. Z. M. Hazil, M. N. Mahdi, M. S. Mohd Azmi, L. K. Cheng, A. Yusof, and A. R. Ahmad, "Software project management using machine learning technique - a review," in *Proceedings of the 2020 8th International Conference on Information Technology and Multimedia (ICIMU)*, pp. 363–370, Selangor, Malaysia, August 2020.
- [8] S. Wang, T. Liu, and L. Tan, "Automatically learning semantic features for defect prediction," *Proceedings of the 38th International Conference on Software Engineering*, vol. 14-22, pp. 297–308, 2016.
- [9] M. Ulan, *Aggregation as Unsupervised Learning in Software Engineering and Beyond*, Linnaeus University Press, Cambridge, MA, USA, 2021.
- [10] N. C. Dang, M. N. Moreno-García, and F. De la Prieta, "Sentiment analysis based on deep learning: a comparative study," *Electronics*, vol. 9, pp. 483–3, 2020.
- [11] J. A. Fadhil, K. T. Wei, and K. S. Na, "Artificial intelligence for software engineering: an initial review on software bug detection and prediction," *Journal of Computer Science*, vol. 16, no. 12, pp. 1709–1717, 2020.
- [12] J. Wen, S. Li, Z. Lin, Y. Hu, and C. Huang, "Systematic literature review of machine learning based software development effort estimation models," *Information and Software Technology*, vol. 54, no. 1, pp. 41–59, 2012.
- [13] Z. Wan, X. Xia, D. Lo, and G. C. Murphy, "How does machine learning change software development practices?" *IEEE Transactions on Software Engineering*, vol. 47, no. 9, pp. 1–1871, 2020.
- [14] A. F. Del Carpio and L. B. Angarita, "Trends in software engineering processes using deep learning: a systematic literature review," in *Proceedings of the 2020 46th Euromicro Conference on Software Engineering and Advanced Applications (SEAA)*, pp. 445–454, Kranj, Slovenia, August 2020.
- [15] F. Meziane and S. Vadera, *Artificial Intelligence in Software Engineering*, Carnegie Mellon University, Pittsburgh, PA, USA, 2010.
- [16] M. Barenkamp, J. Rebstadt, and O. Thomas, "Applications of AI in classical software engineering," *AI Perspect*, vol. 2, no. 1, pp. 1–15, 2020.
- [17] M. Harman, "The role of artificial intelligence in software engineering," in *Proceedings of the 2012 First International Workshop on Realizing AI Synergies in Software Engineering (RAISE)*, pp. 1–6, Zurich, Switzerland, June 2012.
- [18] J. Tate, *Software Process Quality Models: A Comparative Evaluation*, Citeseerx, Pennsylvania, PA, USA, 2003.
- [19] S. Kothawar and R. G. Vajrapu, "Software requirements prioritization practices in software start-ups: a qualitative research based on start-ups in India," vol. 57, 2018.
- [20] W. Guo, D. Mu, J. Xu, P. Su, G. Wang, and X. Xing, "Lemma," in *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security*, pp. 364–379, Toronto, Canada, October 2018.
- [21] M. Shafiq, Z. Tian, A. K. Bashir, X. Du, and M. Guizani, "CorrAUC: a malicious bot-IoT traffic detection method in IoT network using machine learning techniques," *IEEE Internet of Things Journal*, vol. 8, no. 5, pp. 3242–3254, 2021.
- [22] M. Shafiq, Z. Tian, A. K. Bashir, X. Du, and M. Guizani, "IoT malicious traffic identification using wrapper-based feature selection mechanisms," *Computers & Security*, vol. 94, Article ID 101863, 2020.
- [23] M. Shafiq, Z. Tian, A. K. Bashir, A. Jolfaei, and X. Yu, "Data mining and machine learning methods for sustainable smart cities traffic classification: a survey," *Sustainable Cities and Society*, vol. 60, Article ID 102177, September 2020.

## Research Article

# Botnet Attack Detection in IoT Using Machine Learning

**Khalid Alissa** <sup>1</sup>, **Tahir Alyas** <sup>2</sup>, **Kashif Zafar**,<sup>2</sup> **Qaiser Abbas**,<sup>3</sup> **Nadia Tabassum**,<sup>4</sup>  
and **Shadman Sakib** <sup>5</sup>

<sup>1</sup>*Networks and Communications Department, College of Computer Science and Information Technology (CCSIT), Imam Abdulrahman Bin Faisal University (IAU), P.O. Box 1982, Dammam 31441, Saudi Arabia*

<sup>2</sup>*Department of Computer Science, Lahore Garrison University, Lahore 54000, Pakistan*

<sup>3</sup>*Faculty of Computer and Information Systems Islamic University Madinah, Madinah 42351, Saudi Arabia*

<sup>4</sup>*Department of Computer Science, Virtual University of Pakistan, Lahore 54000, Pakistan*

<sup>5</sup>*Department of Finance and Banking, Jahangirnagar University, Bangladesh*

Correspondence should be addressed to Shadman Sakib; [shadman.stu2014@juniv.edu](mailto:shadman.stu2014@juniv.edu)

Received 26 August 2022; Accepted 20 September 2022; Published 4 October 2022

Academic Editor: Ateeq Ur Rehman

Copyright © 2022 Khalid Alissa et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

There are an increasing number of Internet of Things (IoT) devices connected to the network these days, and due to the advancement in technology, the security threats and cyberattacks, such as botnets, are emerging and evolving rapidly with high-risk attacks. These attacks disrupt IoT transition by disrupting networks and services for IoT devices. Many recent studies have proposed ML and DL techniques for detecting and classifying botnet attacks in the IoT environment. This study proposes machine learning methods for classifying binary classes. This purpose is served by using the publicly available dataset UNSW-NB15. This dataset resolved a class imbalance problem using the SMOTE-OverSampling technique. A complete machine learning pipeline was proposed, including exploratory data analysis, which provides detailed insights into the data, followed by preprocessing. During this process, the data passes through six fundamental steps. A decision tree, an XgBoost model, and a logistic regression model are proposed, trained, tested, and evaluated on the dataset. In addition to model accuracy, F1-score, recall, and precision are also considered. Based on all experiments, it is concluded that the decision tree outperformed with 94% test accuracy.

## 1. Introduction

The proliferation of the Internet of Things (IoT) devices has resulted in a steady rise in the volume of IoT-based assaults. One of the most serious IoT risks is the IoT botnet attack, which tries to commit actual, effective, and profitable cybercrimes. IoT botnets are collections of Internet-connected IoT devices that have been infected with malware and are managed remotely by an attacker [1].

The Internet of Things (IoT) systems have significant challenges in offering techniques to detect security vulnerabilities and assaults due to the rapid growth of threats and diversity in attack tactics. As malware is executed, there have been an increasing number of improvements in machine learning/deep learning-based detection tools and techniques that use full-time series data. However, the need to employ full-time series data severely limits existing

works' usefulness [2]. On the other hand, earlier identification would enable better IoT Botnet response proposals. As a result, it reduces the harm caused by possible assaults. The dynamic analysis method examines how malware interacts with its surroundings when it is being executed [3].

These data are extremely important for machine learning and deep learning models detecting malware. The representative approaches required continuous series data collection while the malware is running [4]. In this instance, the malware successfully carried out its goal of information system sabotage and fully exhibited its hostile nature. There are currently available detection techniques for such stages, thus if a DDoS assault performed by an IoT Botnet has already taken place, identifying the DDoS attack, and the IoT Botnet network by itself at this point is not too challenging [5].

The use of bot malware and botnets to support other harmful online activities (such as click fraud, distributed denial of service attacks, and spam and virus distribution). The IoT Botnet lifecycle includes a lengthy scan and propagation phase. If it is feasible to identify and isolate the bots before they launch an actual assault, such as a DDoS, the IoT Botnet detection solution will have a stronger impact [6]. Therefore, it is crucial and required to identify harmful actions of IoT Botnet network components as soon as possible. However, it might be difficult to identify botnets, especially peer-to-peer (P2P) botnets [7]. As a result, we provide a complex traffic reduction strategy in this study that is coupled with a reinforcement learning method. Unquestionably, one of the most fascinating divisions of AI is machine learning. It successfully completes the goal of learning from data with specific machine inputs. It is crucial to understand how ML operates and, consequently, how it might be applied in the future. Training data [8] are entered into the chosen algorithm to begin the machine learning process. The final ML algorithm is developed using the training data, which might be known or unknown data. The method is affected by the type of training data input, and that idea will be discussed in more details shortly [9].

The machine learning algorithm is fed fresh input data to see if it functions properly. Then, the prediction and outcomes are cross-checked [10]. The algorithm is repeatedly retrained if the prediction and results do not line up until the data scientist achieves the desired result. As a result, the ML algorithm is able to continuously train on its own and produce the best solution, steadily improving in accuracy [11].

*1.1. Types of Machine Learning.* The study of machine learning encompasses a wide range of topics and draws inspiration from other domains, including artificial intelligence. The field is centered on learning or gaining abilities or knowledge through practical application. Usually, this entails pulling relevant concepts from the previously collected data. As a result, in the field of machine learning, you may come across a wide variety of learning, ranging from entire fields of study to particular methodologies. Different types of machine learning as shown in Figure 1.

Due to its complexity, machine learning has been separated into two main categories: supervised learning and unsupervised learning, and two ancillary categories: semi-supervised learning and reinforcement learning. Each one has a distinct goal and course of action that produces outcomes and uses different types of data. Supervised learning makes up over 70% of machine learning, whereas unsupervised learning is between 10% and 20%. Semi-supervised and reinforcement learning take up the remaining space.

## 2. Problem Statement

In distributed computing environments, remote access to services has become widespread due to the Internet. Nonetheless, the integrity of data transmission on the

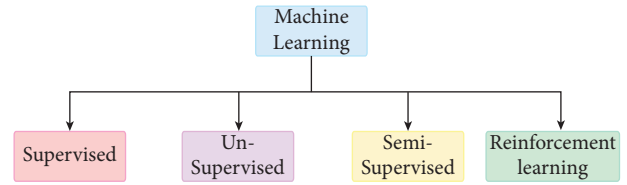


FIGURE 1: Types of Machine learning.

distributed computing platform is hindered by security concerns. Botnets are a prominent threat to Internet security, as are malicious codes. In addition to distributed denial of service (DDoS) attacks, click fraud, phishing, malware distribution, spam emails, and the illegitimate exchange of information or materials, botnets support a wide range of criminal activities. Therefore, developing a robust mechanism for detecting, analyzing, and removing botnets are imperative. At present, botnet detection techniques are reviewed in a variety of ways. However, studies of this type are limited in scopes and lack discussions about the newest botnet detection techniques. The aim of this study is to develop a state-of-the-art machine learning model for botnet detection, utilizing the latest emerging techniques, and analyzing current and past research trends. The study offers a thematic taxonomy for classifying botnet detection techniques and analyses such techniques' implications and critical elements.

## 3. Research Motivation

The field of cybersecurity is always a challenging task for researchers. As a result, cybercriminals constantly research new approaches to identify weaknesses and use them for nefarious and illegal purposes. Malware spreading technique is now growing with new and innovative manners. The malware is then used to carry out further attacks like data exfiltration and denial of service attacks utilizing or on compromised machines.

## 4. Significance of Our Study

Internet of Things (IoT) services and applications have significantly increased due to their functionality and ease of use. Companies have started to develop a variety of Internet of Things (IoT)-based products, ranging from modest personal gadgets like a smartwatch to an entire network of smart grid, smart mining, smart manufacturing, and autonomous driverless vehicles. The overwhelming quantity and ubiquitous presence have enticed potential hackers for data theft and cyberattacks. One of the most significant issues with the Internet of Things is security. This study's main objective is to suggest a novel machine learning algorithm-based model for detecting and thwarting botnet attacks on IoT networks.

## 5. Research Objectives

The objectives of this study are as follows:

- (1) To transform the raw data into machine learning format using data transformation and preprocessing techniques.
- (2) To develop the machine learning model which will be used to classify the botnet attacks.

## 6. Literature Review

The well-organized intrusion representations are used to analyze current and upcoming network outbreaks [12]. Several machine learning algorithms have been established. In this research, a UNSW-NB15 dataset was cast off before the standard KDD99 data set, which depicts current complex attacks and network traffic. An extreme gradient is one of many machine learning algorithms boosting (XGBoost), which delivers extremely efficient and precise data. A subset of the results was chosen. 23 of the 39 useable characteristics were achieved through information gain. Various classifiers such as neural network, multi-logistic regression, nonlinear svm, XGBoost, Naïve Bayes, and random forest are trained and evaluated. From all the XGBoost outperformed with 88% test accuracy, followed by random forest which reported 87.89% accuracy [13].

More experiments were conducted, and the researchers looked up a Deep Neural Network (DNN) for detecting IoT attacks. DNN's ability to correctly identify attacks has been tested on the most commonly used data sets, including KDD-Cup'99, NSL-KDD, and UNSW-NB15. The experimental results demonstrated the precision rate of the projected method using the DNN. It demonstrated that each data set's accuracy rate is greater than 90%. As accuracies reported on the KDD-Cup'99, NSL-KDD, and UNSW-NB15 datasets are 96.30%, 91.50%, and 99.20%, respectively [14].

Similarly, numerous other studies have been conducted in which experts adopt deep learning to detect intrusion. The research includes the deep learning models ANN, DNN, and RNN as an interruption detection system. The dataset UNSW-NB15 was established in diverse files and then categorized into binary classifications with deep learning models to measure abnormal patterns. In this study, the whole dataset was combined in a solo folder for models being tested more fairly than separately for a separate file. The dataset attack families were then used as new labels, resulting in a multi-classification labelled dataset. The improved dataset categorized deep learning performance into dual arrangement groups (Binary and Multi-Class). The deep learning models that we propose show the accuracy in multi-class classification was 99.59%, and the accuracy in binary classification was 99.26% [15]. The comparison between research shows the competence of DL and ML representations in the improved dataset using accuracy and loss.

The paper proposed and tested the IGRFRFE fusion collection technique on behalf of MLP incursion concealment systems on the UNSW-NB15 modern IDS set. IGRFRFE is made up of two feature reduction steps, one of which is IGRF recursive feature elimination with MLP and ensemble feature selection. The screen option collection technique was used to reduce the feature subset search space, which is a mixture of IG and RF Importance. Then, as a

wrapper feature collection method, recursive feature elimination (RFE) was used to eliminate terminated features on the concentrated feature subsets. The effects indicate that the option measurement is concentrated from 42 to 23, while the MLP precision is upgraded from 82.25% to 84.24%. The outcomes on the UNSW-NB15 dataset authorize that the projected process can expand irregularity, and detection accuracy while dropping feature measurement [16].

ML methods can detect data based on prior experience and distinguish between normal and abnormal data. The CNN DL technique was established in research work conducted in 2021 to resolve the difficulties of identifying network intrusion. The CNN algorithm was accomplished using the UNSW NB15. In general, the data covers binary types for usual and attack data. The tentative results verified that the anticipated model accomplishes extreme detection accuracy of 93.5%, and also evaluation metrics were used to amount to the performance of the CNN algorithm [17].

According to the tentative results, the original KDD99 features are less effective than the KDD99 data set's simulated UNSW-NB15 features. However, when datasets are compared, the precision of the KDD99 dataset is higher than that of the UNSW-NB 15. The FAR of the KDD99 is lesser than that of the UNSWNB 15 dataset. However, the reported accuracy of the proposed model is 98.89% [18].

Recent improvements in machine learning consumed a preferred tool for various classification and analytical difficulties. The analysis provides information, investigates challenges, fundamental analyses of data in terms of security, and forecasts future opportunities for machine learning in networking. From all the proposed classifiers, the random forest outperformed with 86.99% accuracy while Ada boost performed the least with 83.67% test accuracy [19].

## 7. Solution Design and Implementation

*7.1. Conceptual Description of the Solution.* The methodology proposed for this research has strictly followed the classical machine learning pipeline. Following all steps, the data was passed to the proposed classifiers and evaluated. The flow-chart of modelling can be illustrated in Figure 2.

Whenever a dataset project is undertaken, the first phase involves the collection of datasets. For this research, the well-known dataset known as UNSW-NB15 has been collected from Kaggle. The dataset was then passed through the exploratory data analysis (EDA) phase to perform the statistical analysis and get meaningful insights into dataset attributes [20].

The proposed methodology consists of feather extraction, traffic reduction, and a multi-layer network classifier to detect the botnet from the normal traffic. In the first phase, traffic is reduced by filtering the TCP control packet after the feather extraction. After extracting the features, the model is trained for botnet detection and legitimate traffic, as shown in Figure 3.

*7.2. Dataset Description.* In this research, the models are built and trained to classify the botnet attacks. For this purpose, a publicly available dataset known as UNSW-NB15

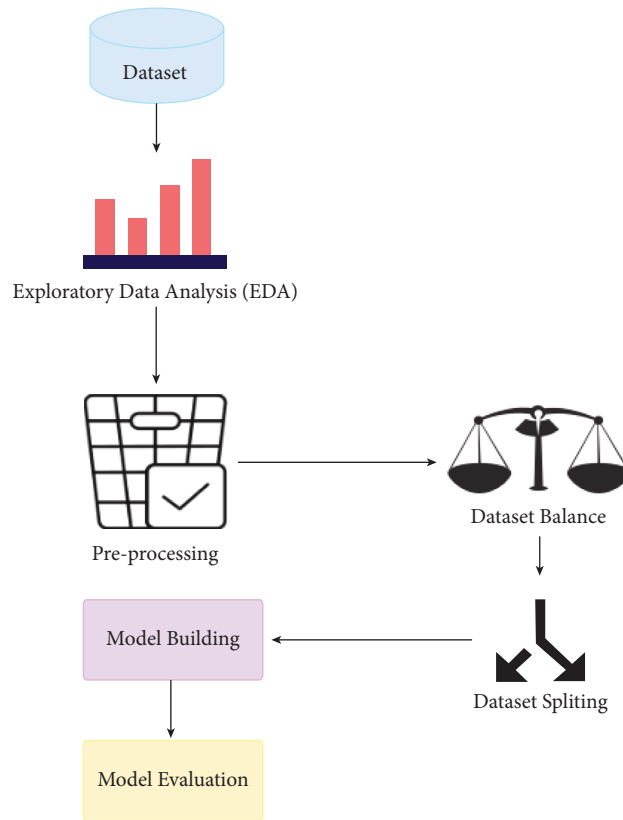


FIGURE 2: Flow chart for modelling.

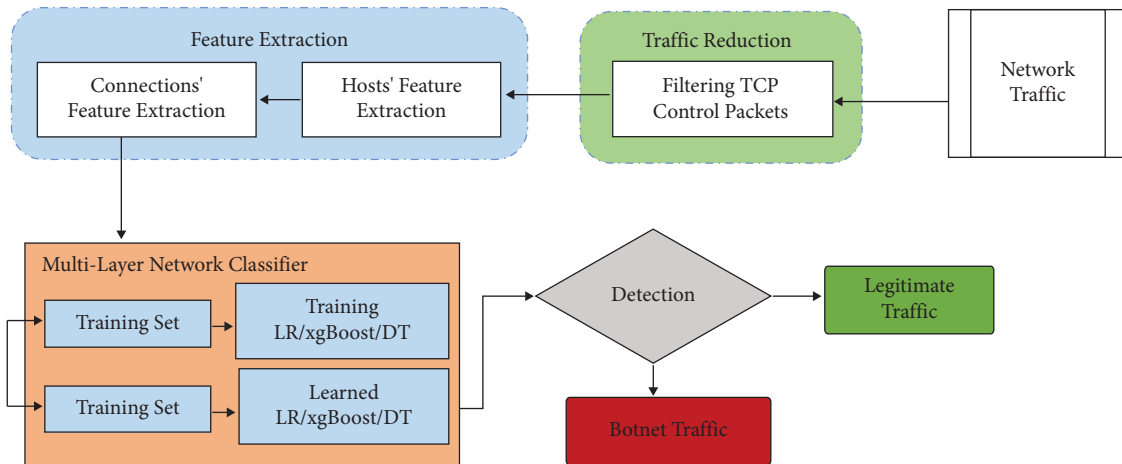


FIGURE 3: Proposed methodology.

is collected from Kaggle [21]. The dataset is published by the IXIA Perfect Storm tool. The Australian Centre for Cyber Security (ACCS) is an authentic botnet classification dataset. The UNSW-NB 15 dataset was created by the IXIA PerfectStorm tool in the Cyber Range Lab of the Australian Centre for Cyber Security (ACCS) to generate a hybrid of real modern normal activities and recent synthetic attack behaviors [22].

This dataset has nine attacks labelled, such as Fuzzers, Analysis, Backdoors, DoS, Exploits, Generic, Reconnaissance, Shellcode, and Worms [23–25]. Using the tools

Argus, and Bro-IDS, 12 algorithms have been developed to generate a total of 49 functions in class labels. These features are described in the UNSW-NB15\_features.csv file. The total number of records is 2 million and 540,044 are stored in four CSV files UNSW-NB151.csv, UNSW-NB152.csv, UNSW-NB153.csv, and UNSW-NB154.csv. The name of the ground truth table is UNSW-NB15GT.csv and the name of the list of event files are UNSW-NB15LIST\_EVENTS.csv.

The partitions for this dataset are configured as training sets and test sets, namely UNSWNB15training-set.csv and UNSWNB15testing-set.csv, respectively. The training set has

175,341 records, and the test set has various types of attacks and regular 82,332 records.

The dataset has 45 attributes. Details of attributes are described in Table 1.

This is a binary class classification dataset as its label column has just two values (0 and 1). 1 represents that it is the attack record else 0 in the case of a normal record. In total, the dataset has 30 integers, 11 floats, and 4 categorical attributes.

The acute process of performing an initial investigation on data to discover patterns, spot noise, and outliers, to test a hypo study, and to check assumptions with the help of summary statistics and graphical representation is called exploratory data analysis and is commonly known as EDA. When analyzing a dataset, it is important to do both statistical and graphical analyses.

The visualization reveals that this is the binary class classification dataset as it has only 2 classes (0 and 1). Secondly, the dataset is highly imbalanced as the 0 class is almost half of 1. The count of class 1 and 0 is 164673 and 93000, respectively as shown in Figure 4. If this thing remains unhandled then the model will not be trained accurately, affecting the model performance and leading to the miss classification.

This is the frequency chart of the attack cat column, representing 10 different categories of attacks, out of which the normal is the highest and the worm's category has the minimum frequency as shown in Figure 5. Other categories are generic, exploits, fuzzers, DoS, surveillance, analysis, backdoor, and shellcode. The proto is an attribute which enlists all the transaction protocols. It has 10 unique values, including tcp, udp, ospf, gre protocols, and the percentage of each protocol is present. Tcp has the highest frequency, followed by udp where ipv6 occurred the most least.

Figure 6 and Figure 7 are illustrations of service and state protocol, respectively. The service attribute has 11 unique values, of which 7 attributes' frequency almost equals to zero, whereas the state attribute has 13 unique values, around which 5 attributes have a frequency which almost equals to zero. Skewness is a measure of asymmetry or distortion of symmetric distribution. It measures the deviation of the given distribution of a random variable from a symmetric distribution, such as normal distribution. A normal distribution is without any skewness, as it is symmetrical on both sides. Hence, a curve is regarded as skewed if it is shifted towards the right or the left. The skewness of all numerical features of this research is shown in Figure 8.

All the numerical attributes have normal skewness and most of them are either left or right skewed. Only the id column has normal distribution because it has all unique values.

Values range from 0 to 60. Most of the values are close to 0 and lesser than 20. ct\_dst\_ltm highly corr with ct\_dst\_sport\_ltm ct\_src\_ltm corr with ct\_src\_dport\_ltm ct\_src\_dport\_ltm corr with ct\_dst\_sport\_ltm.

Numerical feature with a small discrete set of values. Normal data has 0 as most of its values. Anomaly has most of its value 2. There are few attacks with a value 1 and very little nonattack with 1.

This feature has 55 unique values. Attack data has value 0 most of the time, but that is also way too little compared with nonattack data. Most of the nonattack data have value of 1 and very few 2,3,4, 19,21,23 but very few in number not visible in the graph a.

Figure 9 has a range up to 800,000. Normal data is distributed over a very wide range up to 200,000. For attack data, there is a huge peak close to 0, and distribution of values is very narrow, whereas the Dload feature has a high correlation with target feat, the 0.35 Feature has a huge range of values up to 1e8. We can visualize better in a log scale. Normal data are distributed all over, has values close to 0 and very large values; for attack data, all the values are very close to 0. IN log scale we can see that values are between 3 and 15.

Most of the values for nonattack data are 29. There are some 0 and very few 252, as shown in Figure 10. There are lots of 0 in attack data, no of 0 in attack is more than nonattack, and 252, which is also higher than nonattack. Boxplots are used to visualize the outliers/noise present in data. The most important attributes have been visualized using a boxplot.

These are the boxplots of attributes labels, is\_sm\_ips\_ports, ct\_src\_ltm, response\_body\_len, swin, dttl and sttl as shown in Figure 11.

Figure 12 illustrates the attributes attack cat and no of events. The second and most important step of any machine learning project is dataset processing. Hence the dataset is in raw format and cleaned it, and transformed the data into the form acceptable by machine learning classifiers.

*7.3. Data Pre-processing.* The data mining technique used to convert raw data into valuable information for machines is known as preprocessing. The following steps have been performed in preprocessing. It is observed that the real-world data is often incomplete, is inconsistent, and contains a lot of errors.

As illustrated in Figure 13, datasets are preprocessed by using the given fundamental steps.

*7.3.1. Handle Null/Missing Values and Duplicate Data.* Fortunately, the dataset has no null or missing values, as shown in Figure 14. Moreover, there is no redundant data too. Till this point, the dataset is clean, but *i* need further preprocessing such as label encoding and feature extraction.

*7.4. Label Encoding.* The dataset has four categorical columns that need encoding to feed into the model. To transform it into a numeric form, the label encoder is used.

Figure 15 and Figure 16 shows the before and after of encoding, and all the data has been transformed successfully.

*7.5. Feature Engineering.* One of the most crucial steps of preprocessing is a selection of features. There are many ways of feature selection. The most suitable features have been extracted and selected for this research using the correlation technique.

TABLE 1: Few dataset description attribute.

Sr#	Attribute	Description	Data Type
1	Id	The unique serial number of records	Integer
2	Dur	Record total duration	Float
3	Proto	Transaction protocol	Object
4	Service	Http, ftp, smtp, ssh, dns, ftp-data, irc and (-) if not much used service	Object
5	State	Indicates to the state and its dependent protocol	Object
6	Spkts	Source to destination packet count	Integer
7	Dpkts	Destination to source packet count	Integer
8	Sbytes	Source to destination transaction bytes	Integer
9	Dbytes	Destination to source transaction bytes	Integer

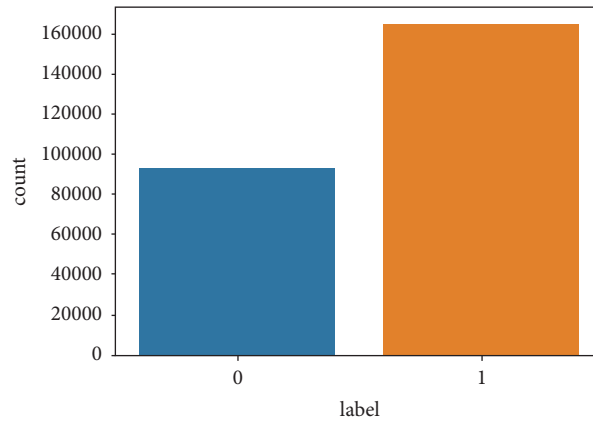


FIGURE 4: Attribute label plot.

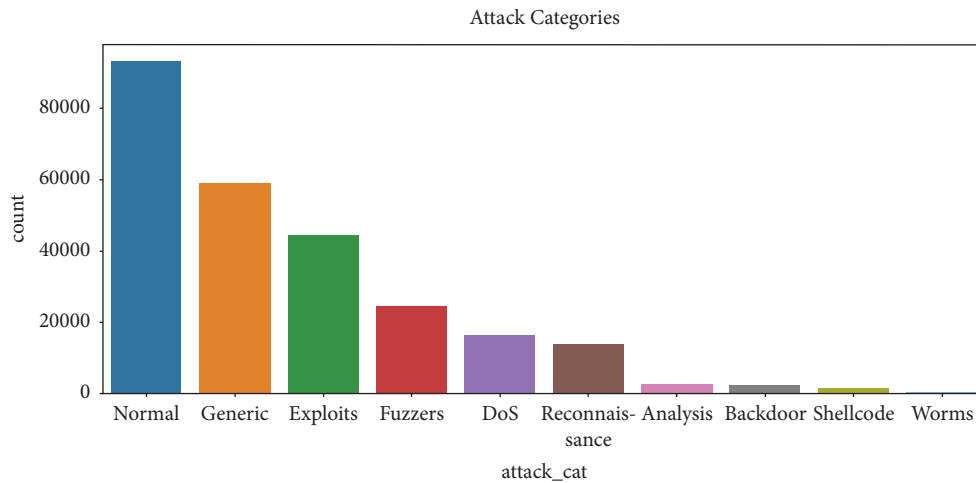


FIGURE 5: Attribute "attack\_cat" plot.

Three types of correlation exist between the features: neutral, positive, and negative. Figure 17 illustrates the heatmap of correlation that exists between the features. With 0.4 threshold the discarded features are 'id', 'sloss', 'dloss', 'dpkts', 'dwin', 'time', 'ct\_srv\_dst', 'ct\_src\_dport\_ltm', 'ct\_dst\_src\_ltm'. These features negatively correlate with label and will affect the model negatively.

**7.6. Balance Dataset.** Imbalanced data typically refers to a problem with classification problems where the classes are not represented equally. Most classification data sets do not have an exactly equal number of instances in each class, but a slight difference often does not matter.

Specialized techniques which can be used to balance the dataset are



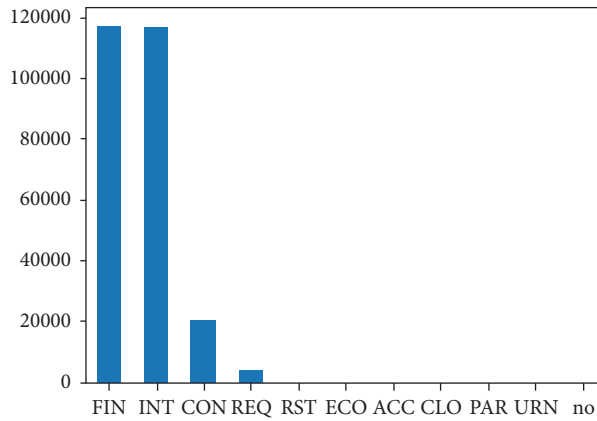


FIGURE 6: Attribute "state" description.

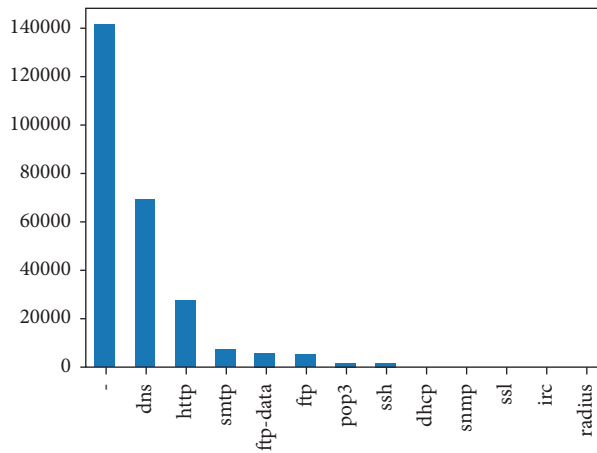


FIGURE 7: Attribute "service" description.

- (1) Under-sampling
- (2) Oversampling
- (3) SMOTE

In this research SMOTE oversampling has been used. SMOTE is a technique generally known as Synthetic Minority Oversampling Technique. This is a systematic algorithm used to generate synthetic samples. As the name implies, SMOTE is an oversampling method. Instead of making a copy, it works by making a synthetic sample from a subclass. The algorithm selects two or more similar instances (using a distance measure). It perturbs one instance with one attribute at a time, in random quantities within the range of differences from adjacent instances. After applying the SMOTE oversampling technique, the dataset is now balanced, and the final shape of the dataset is (329346,37) instead of (257673,36). Now the dataset is cleaned, transformed, and balanced. It is ready to train a machine learning model.

## 8. Result and Analysis

Three machine learning classifiers, Decision Tree, XgBoost, and Logistic Regression, are trained, tested, evaluated, and

compared in this dataset. The comparative analysis is also performed for both unbalanced and balanced datasets.

**8.1. Decision Tree.** Firstly, the dataset was split into train and test by using the train test split method. 80% was used for training, while 20% was used for testing. This model was trained for both balanced and unbalanced datasets, and the following results have been achieved in Table 2.

From Figure 18 and Figure 19 we can see that all the evaluation parameters (precision, recall, and f1-score) are 100, which indicates that there is a problem in learning. The model is only learning one class and miss classifying other.

These are the confusion matrix in which there is no true negative and false positive for an unbalanced dataset present in Figures 20 and 21.

**8.2. XgBoost.** XgBoost is a boosting classifier with 80% training and 20% testing. The model was trained for both balanced and unbalanced datasets, and the following results have been achieved as shown in Table 3.

From Figures 22 and 23 we can see that all the evaluation parameters (precision, recall, and f1-score) are 100,

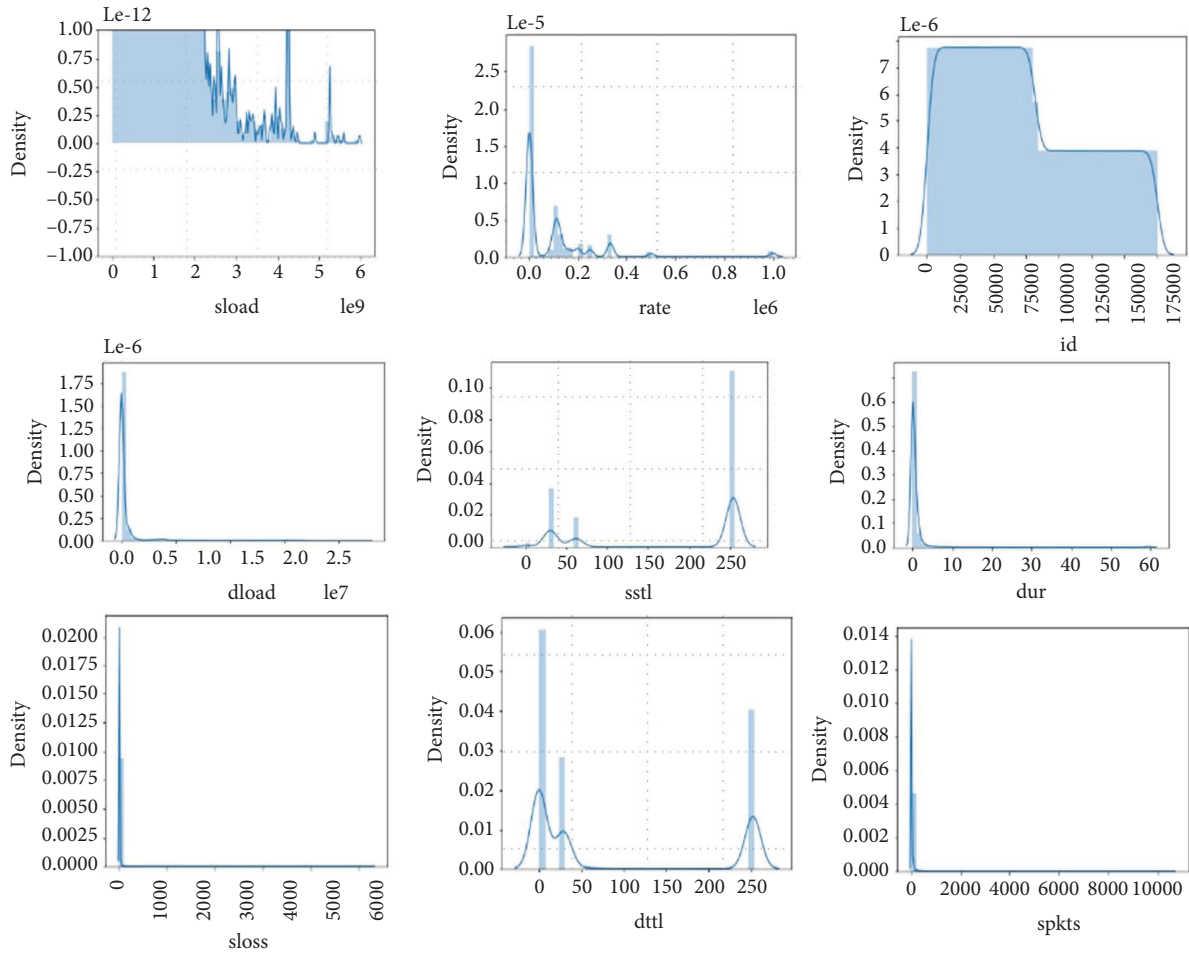


FIGURE 8: Skewness plots.

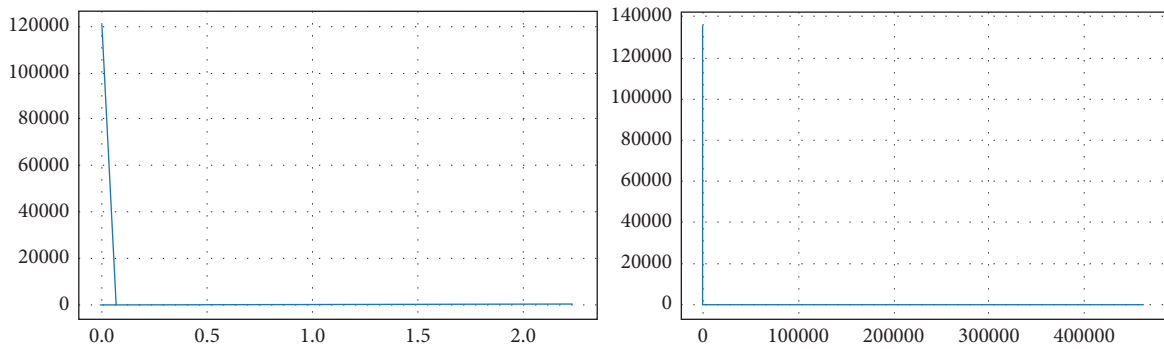


FIGURE 9: Pair plots of attributes “djit and dload”.

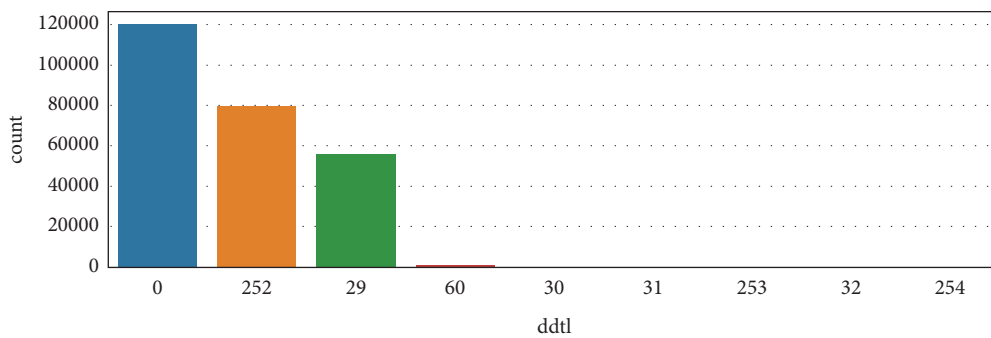


FIGURE 10: Attribute “ddtl” description.

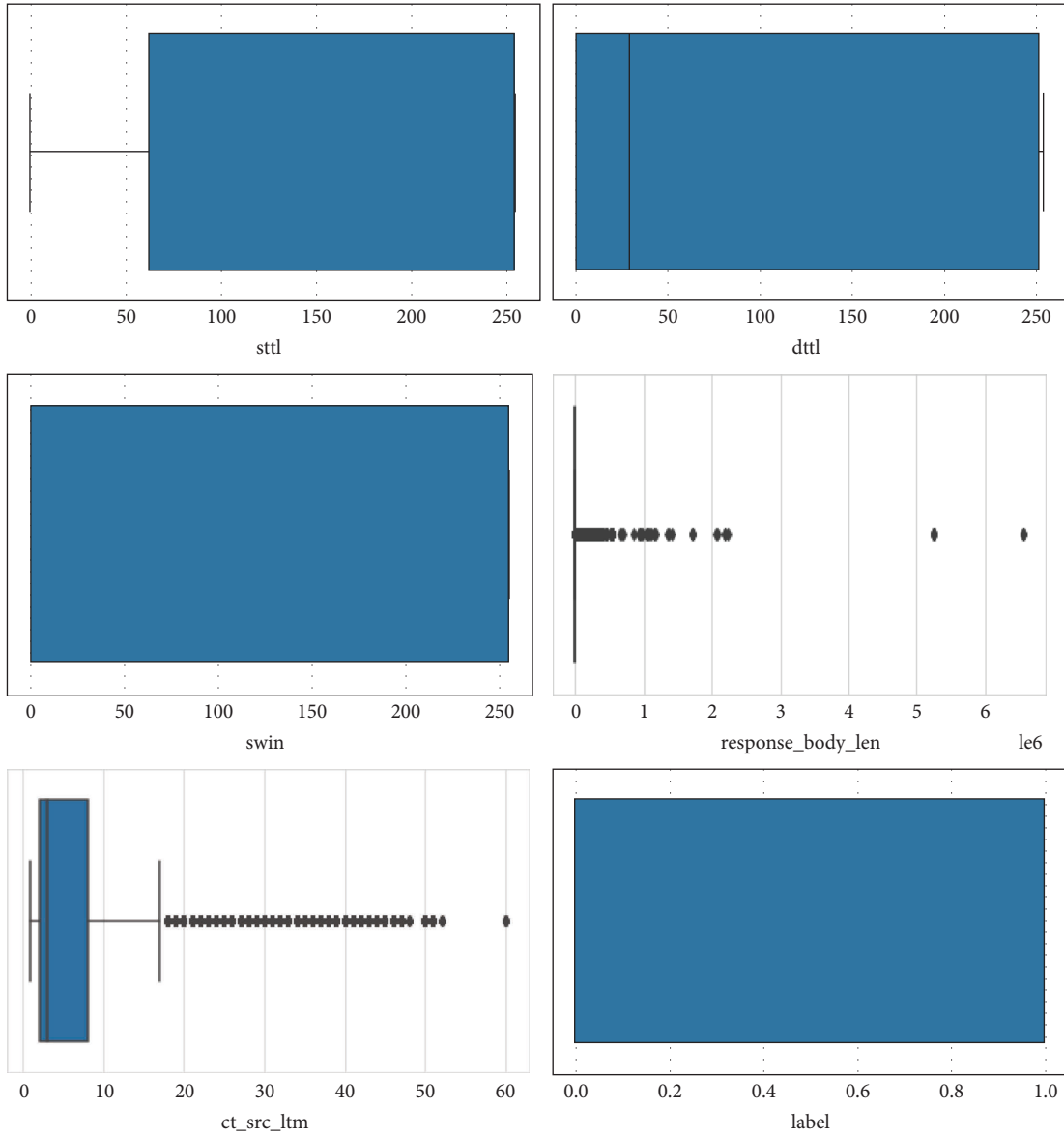


FIGURE 11: Box plots of all numeric attributes.

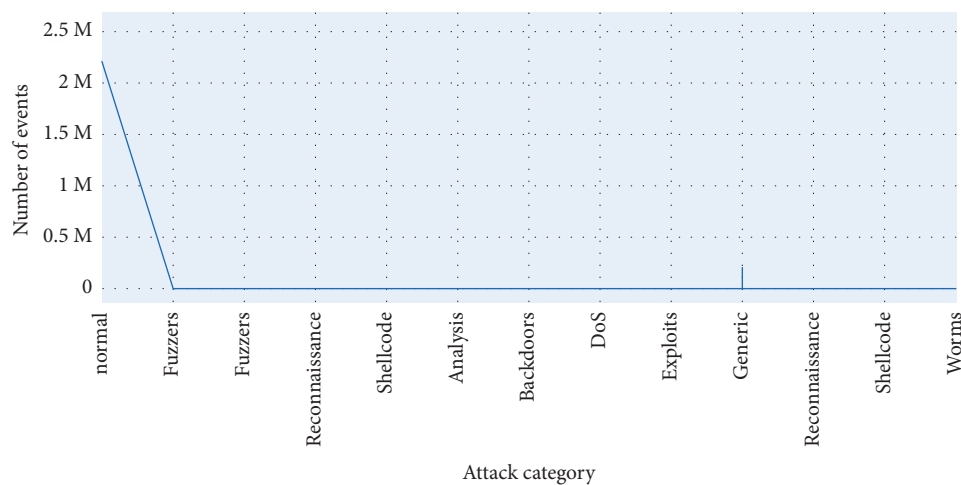


FIGURE 12: Comparison plot of attributes “attack\_cat VS no. of events”.

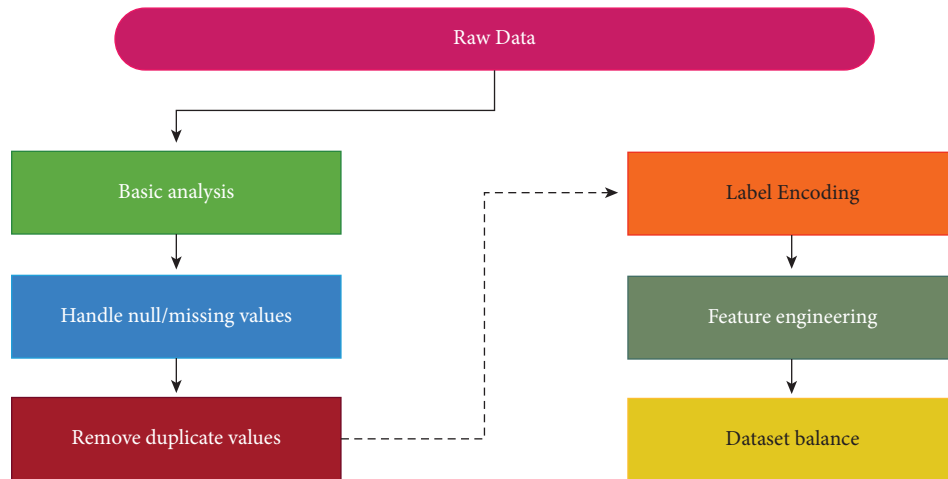


FIGURE 13: Pre-processing steps.

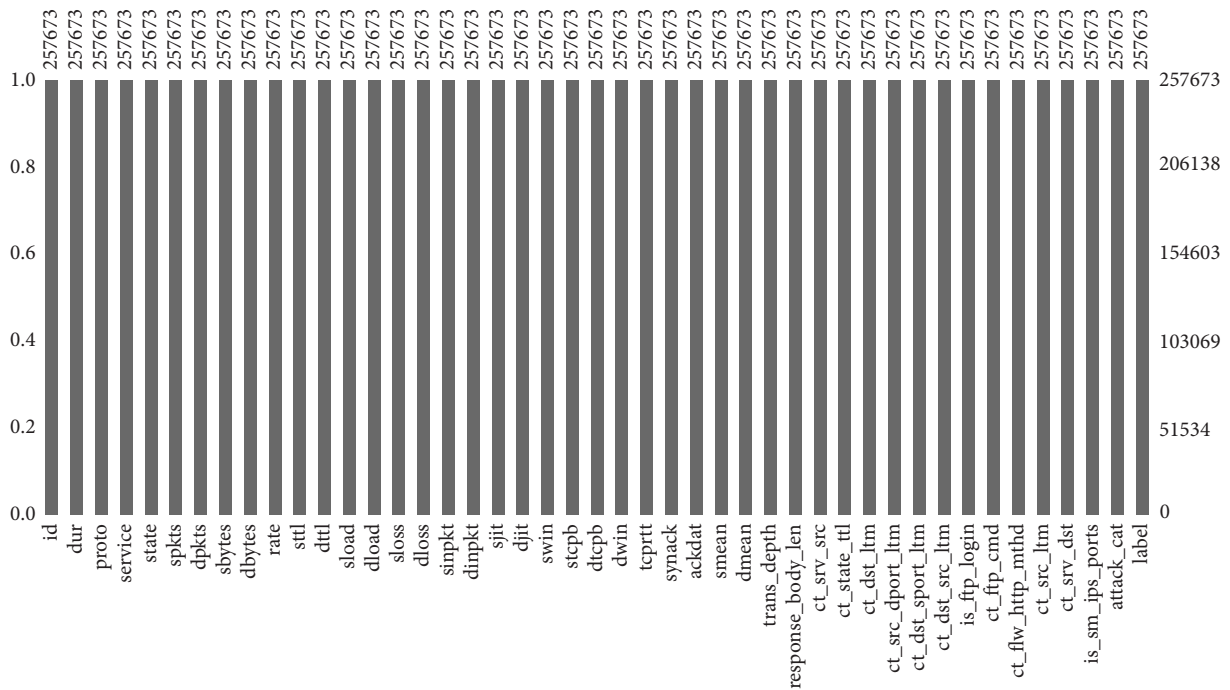


FIGURE 14: Plot of missing/null values.

	id	dur	proto	service	state	spkts	dpkts	sbytes	dbytes	rate	...	ct_dst_sport_ltm	ct_dst_src_ltm	is_ftp_login	ct_ftp_cmd
0	1	0.000011	udp	nil	INT	2	0	496	0	90909.0902	...	1	2	0	C
1	2	0.000008	udp	nil	INT	2	0	1762	0	125000.0003	...	1	2	0	C
2	3	0.000005	udp	nil	INT	2	0	1068	0	200000.0051	...	1	3	0	C
3	4	0.000006	udp	nil	INT	2	0	900	0	166666.6608	...	1	3	0	C
4	5	0.000010	udp	nil	INT	2	0	2126	0	100000.0025	...	1	3	0	C

FIGURE 15: Before encoding.

	id	dur	proto	service	state	spkts	dpkts	sbytes	dbytes	rate	...	ct_dst_sport_ltm	ct_dst_src_ltm	is_ftp_login	ct_ftp_cmd
0	1	0.000011	119	6	5	2	0	496	0	90909.0902	...	1	2	0	C
1	2	0.000008	119	6	5	2	0	1762	0	125000.0003	...	1	2	0	C
2	3	0.000005	119	6	5	2	0	1068	0	200000.0051	...	1	3	0	C
3	4	0.000006	119	6	5	2	0	900	0	166666.6608	...	1	3	0	C
4	5	0.000010	119	6	5	2	0	2126	0	100000.0025	...	1	3	0	C

FIGURE 16: After encoding.

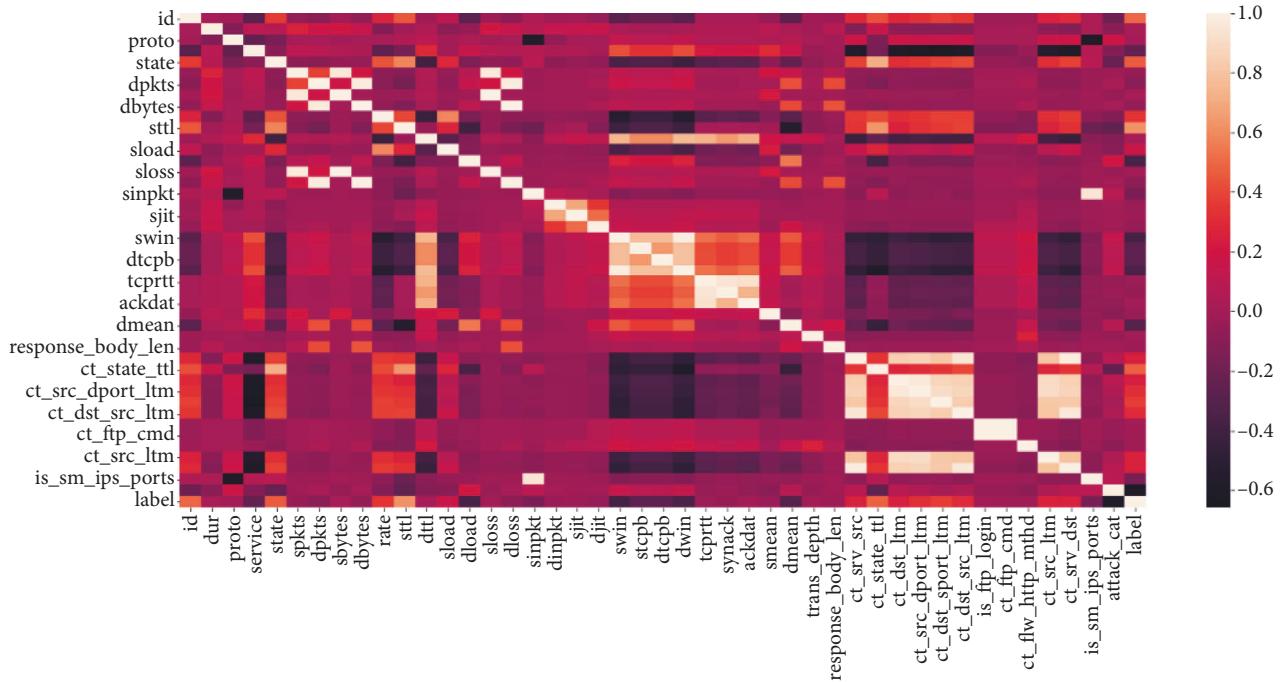


FIGURE 17: Heapmap - feature engineering.

	precision	recall	f1-score	support
0	1.00	1.00	1.00	18675
1	1.00	1.00	1.00	32860
accuracy			1.00	51535
macro avg	1.00	1.00	1.00	51535
weighted avg	1.00	1.00	1.00	51535

FIGURE 18: Classification report of imbalanced data - DT.

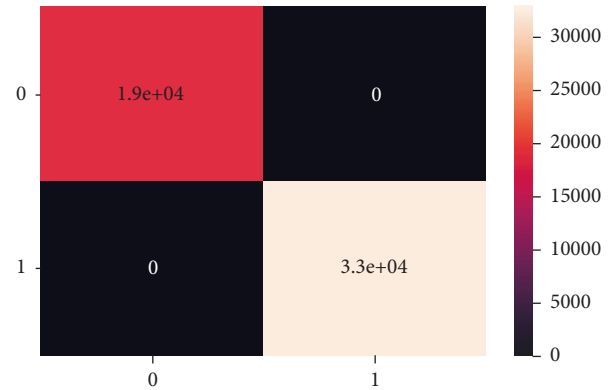


FIGURE 20: Confusion matrix of imbalanced data - DT.

TABLE 2: Accuracy table - Decision tree.

Sr#	Dataset	Train accuracy (%)	Test accuracy (%)
1	Un-balanced	100	100
2	Balanced	100	94

The classification report for both cases is.

	precision	recall	f1-score	support
0	0.94	0.94	0.94	32977
1	0.94	0.94	0.94	32893
accuracy			0.94	65870
macro avg	0.94	0.94	0.94	65870
weighted avg	0.94	0.94	0.94	65870

FIGURE 19: Classification report of balanced data - DT.

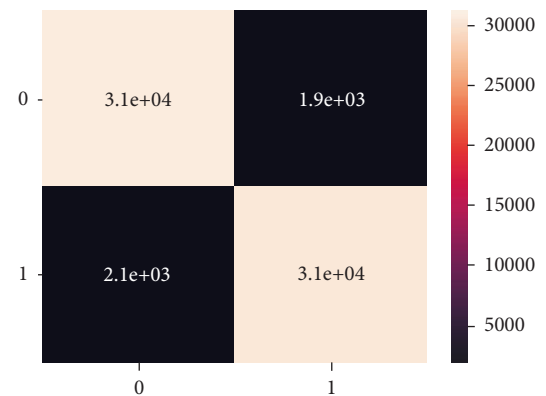


FIGURE 21: Classification report of balanced data - DT.

TABLE 3: Accuracy table - XgBoost.

Sr#	Dataset	Train accuracy (%)	Test accuracy (%)
1	Imbalanced	100	100
2	Balanced	100	93

	precision	recall	f1-score	support
0	1.00	1.00	1.00	18675
1	1.00	1.00	1.00	32860
accuracy			1.00	51535
macro avg	1.00	1.00	1.00	51535
weighted avg	1.00	1.00	1.00	51535

FIGURE 22: Classification report of imbalanced data - XgBoost.

	precision	recall	f1-score	support
0	0.94	0.94	0.94	32977
1	0.94	0.94	0.94	32893
accuracy			0.94	65870
macro avg	0.94	0.94	0.94	65870
weighted avg	0.94	0.94	0.94	65870

FIGURE 23: Classification report of balanced data-XgBoost.

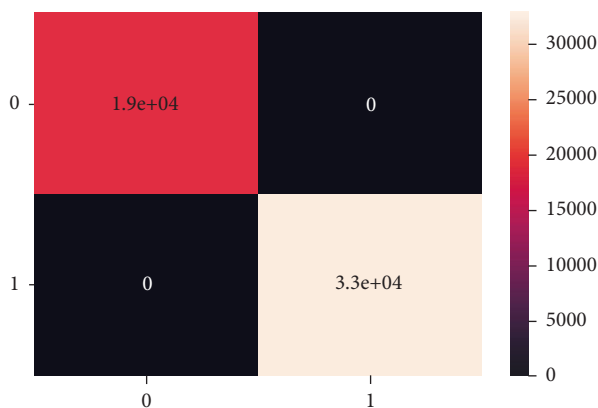


FIGURE 24: Confusion matrix of imbalanced data-XgBoost.

indicating a problem in learning, and the model is only learning the one class and miss classifying others.

Unlike the balanced dataset confusion matrix, there is no classification in the unbalanced dataset confusion matrix as shown in Figures 24 and 25.

**8.3. Logistic Regression.** Logistic regression is the regression technique used to classify binary data. In this 80% training

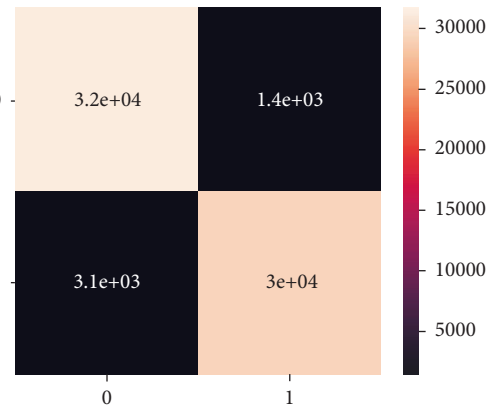


FIGURE 25: Confusion matrix of balanced data-XgBoost.

TABLE 4: Accuracy table - Logistic regression.

Sr#	Dataset	Train accuracy (%)	Test accuracy (%)
1	Unbalanced	100	100
2	Balanced	100	78

	precision	recall	f1-score	support
0	1.00	1.00	1.00	18675
1	1.00	1.00	1.00	32860
accuracy			1.00	51535
macro avg	1.00	1.00	1.00	51535
weighted avg	1.00	1.00	1.00	51535

FIGURE 26: Classification report of imbalanced data - LR.

	precision	recall	f1-score	support
0	0.73	0.91	0.81	32977
1	0.88	0.66	0.75	32893
accuracy			0.78	65870
macro avg	0.80	0.78	0.78	65870
weighted avg	0.80	0.78	0.78	65870

FIGURE 27: Classification report of balanced data - LR.

and 20% testing split dataset has been passed to the model for training of both balanced and unbalanced dataset, and the following results have been achieved as shown in Table 4.

From Figures 26 & 27, we can see that all the evaluation parameters (precision, recall, and f1-score) are 100, indicating a problem in learning. The model is only learning one class and miss classifying others. The performance of logistic regression was not good compared to the decision tree and XgBoost.

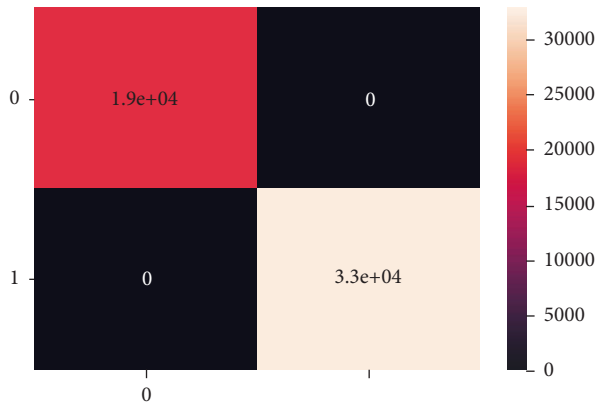


FIGURE 28: Confusion matrix of imbalanced data, LR.

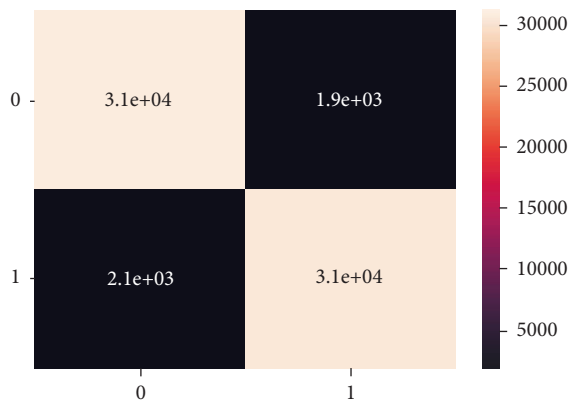


FIGURE 29: Confusion matrix of balanced data, LR.

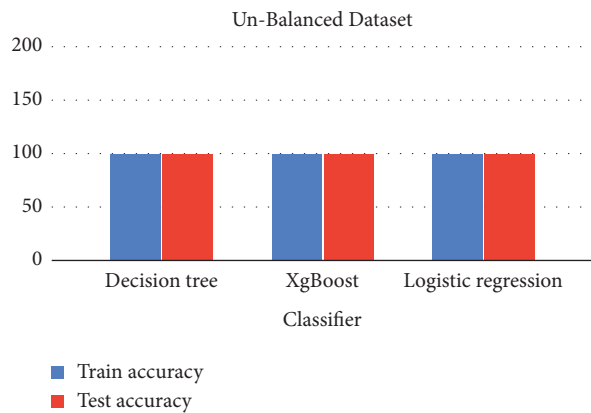


FIGURE 30: Comparative analysis of unbalanced data.

From all the trained models, the decision tree outperformed with a slightly higher accuracy than the decision tree by predicting a few true positive values and logistic regression has the least performance.

These confusion matrix results are the same for the unbalanced dataset as there is no true negative and false positive present unlike the balanced dataset confusion matrix, there is no classification there, as shown in Figures 28 and 29.

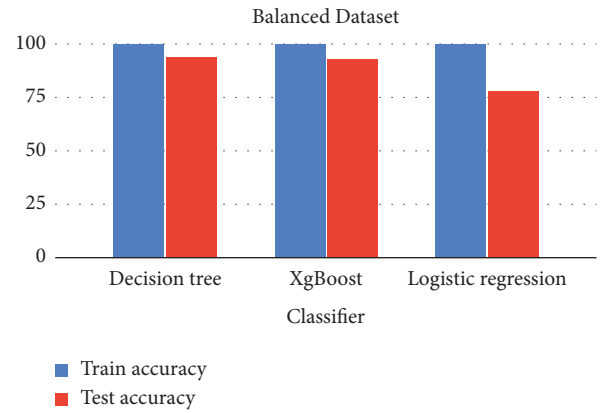


FIGURE 31: Comparative analysis of balanced data.

The charts summarized all the results and experiments of this research. With a balanced dataset, the decision tree outperformed with 94% accuracy, as shown in Figures 30 and 31.

### 9. Conclusion

Cyber-attacks involving botnets are multi-stage attacks and primarily occur in IoT environments; they begin with scanning activity and conclude with distributed denial of service (DDoS). Most existing studies concern detecting botnet attacks after IoT devices become compromised and start performing DDoS attacks. Furthermore, most machine learning-based botnet detection models are limited to a specific dataset on which they are trained. Consequently, these solutions do not perform well on other datasets due to the diversity of attack patterns. In this work, UNSW-NB15, the most generalized dataset publicly available, is used. EDA (Exploratory Data Analysis) is the statistical analysis phase through which the whole dataset is analyzed.

The model will be able to be trained on a large data set in the future. Machine learning classifiers such as Random Forest and SVM can also be tested. As well as ResNet50 and LSTM models, deep learning models can also be used in runtime Botnet detection. Besides being integrated with front-end web applications, the research' model can also be used with back-end web applications.

### Data Availability

The data used in this paper can be requested from the corresponding author upon request. Data will be provided on request.

### Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this work.

### References

[1] A. Kumar, A. K. Singh, I. Ahmad et al., "A novel decentralized blockchain architecture for the preservation of privacy and data security against cyberattacks in healthcare," *Sensors*, vol. 22, no. 15, pp. 1-14, 2022.

- [2] T. Alyas, I. Javed, A. Namoun, A. Tufail, S. Alshmrany, and N. Tabassum, "Live migration of virtual machines using a mamdani fuzzy inference system," *Computers, Materials & Continua*, vol. 71, no. 2, pp. 3019–3033, 2022.
- [3] M. S. Mazhar, Y. Saleem, A. Almogren et al., "Forensic analysis on internet of things (IoT) device using machine to machine (M2M) framework," *Electronics*, vol. 11, no. 7, p. 1126, 2022.
- [4] T. Alyas, K. Alissa, M. Alqahtani et al., "Multi-Cloud integration security framework using honeypots," *Mobile Information Systems*, vol. 2022, pp. 1–13, Article ID 2600712, 2022.
- [5] T. Kalsoom, N. Ramzan, S. Ahmed, and M. Ur-Rehman, "Advances in sensor technologies in the era of smart factory and industry 4.0," *Sensors*, vol. 20, p. 6783, 2020.
- [6] M. Ahmad, T. M Ghazal, and N. Aziz, "A survey on animal identification techniques past and present," *International Journal of Computational and Innovative Sciences*, vol. 1, no. 2, pp. 1–7, 2022.
- [7] S. T. Reddy and G. K. Shyam, "A machine learning based attack detection and mitigation using a secure SaaS framework," *Journal of King Saud University – Computer and Information Sciences*, vol. 34, no. 7, pp. 4047–4061, 2022.
- [8] N. Tabassum, T. Alyas, M. Hamid, M. Saleem, S. Malik, and S. Binish Zahra, "Qos based cloud security evaluation using neuro fuzzy model," *Computers, Materials & Continua*, vol. 70, no. 1, pp. 1127–1140, 2022.
- [9] A. M. Araujo, A. Bergamini de Neira, and M. Nogueira, "Autonomous machine learning for early bot detection in the internet of things," *Digital Communication Networks*, in press, 2022.
- [10] S. Muneer and M. A Rasool, "Systematic review: explainable Artificial Intelligence (XAI) based disease prediction," *International Journal of Advanced Sciences and Computing*, vol. 1, no. 1, pp. 1–6, 2022.
- [11] R. Abrantes, P. Mestre, and A. Cunha, "Exploring dataset manipulation via machine learning for botnet traffic," *Procedia Computer Science*, vol. 196, pp. 133–141, 2022.
- [12] A. Kumar, M. Shridhar, S. Swaminathan, and T. J. Lim, "Machine learning-based early detection of IoT botnets using network-edge traffic," *Computers & Security*, vol. 117, Article ID 102693, 2022.
- [13] A. Husain, A. Salem, C. Jim, and G. Dimitoglou, "Development of an efficient network intrusion detection model using extreme gradient boosting (XGBoost) on the UNSW-NB15 dataset," in *Proceedings of the 2019 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*, pp. 1–7, Ajman, UAE, December 2019.
- [14] A. Alharbi and K. Alsubhi, "Botnet detection approach using graph-based machine learning," *IEEE Access*, vol. 9, pp. 99166–99180, 2021.
- [15] A. U. Rehman, R. A. Naqvi, A. Rehman, A. Paul, M. T. Sadiq, and D. Hussain, "A trustworthy SIoT aware mechanism as an enabler for citizen services in smart cities," *Electronics*, vol. 9, no. 6, p. 918, 2020.
- [16] G. Mahalakshmi, E. Uma, and A. M. Vinitha, "Advances in parallel computing technologies and applications," *Intrusion Detection System Using Convolutional Neural Network on UNSW NB15 Dataset, Advances in Parallel Computing*, vol. 40, pp. 1–8, 2021.
- [17] T. Kalsoom, S. Ahmed, P. M. Rafi-ul-Shan et al., "Impact of IoT on Manufacturing Industry 4.0: a new triangular systematic review," *Sustainability*, vol. 13, no. 22, p. 12506, 2021.
- [18] S. Choudharya and N. Kesswanib, "Analysis of KDD-cup'99, NSL-KDD and UNSW-NB15 datasets using deep learning in IoT," in *Proceedings of the International Conference on Computational Intelligence and Data Science*, Rajasthan, India, May 2019.
- [19] S. Bagui, M. Walauski, R. DeRush, H. Praviset, and S. Boucugnani, "Spark configurations to optimize decision tree classification on UNSW-NB15," *Big Data and Cognitive Computing*, vol. 6, no. 2, p. 38, 2022.
- [20] N. Iqbal, S. Abbas, M. A. Khan, T. Alyas, A. Fatima, and A. Ahmad, "An RGB image cipher using chaotic systems, 15-puzzle problem and DNA computing," *IEEE Access*, vol. 7, pp. 174051–174071, 2019.
- [21] P. Jacob, "UNSW-NB15 dataset feature selection and network intrusion detection using deep learning," *International Journal of Recent Technology and Engineering*, vol. 7, no. 5S2, 2019.
- [22] Z. A. Chughtia, M. Awais, and A. Rasheed, "Distributed autonomous organization security in blockchain:(DAO attack)," *International Journal of Computational and Innovative Sciences*, vol. 1, no. 2, pp. 51–68, 2022.
- [23] Z. Shao, S. Yuan, and Y. Wang, "Adaptive online learning for IoT botnet detection," *Information Sciences*, vol. 574, pp. 84–95, 2021.
- [24] S. Kanwal, F. Tao, A. Almogren, A. Ur Rehman, R. Taj, and A. Radwan, "A robust data hiding reversible technique for improving the security in e-health care system," *Computer Modeling in Engineering and Sciences*, vol. 134, no. 1, pp. 201–219, 2023.
- [25] M. Alauthman, N. Aslam, M. Al-kasassbeh, S. Khan, A. Al-Qerem, and K. K. Raymond Choo, "An efficient reinforcement learning-based Botnet detection approach," *Journal of Network and Computer Applications*, vol. 150, Article ID 102479, 2020.



## Retraction

# Retracted: Artificial Bee Colony Based Gabor Parameters Optimizer (ABC-GPO) for Modulation Classification

### Computational Intelligence and Neuroscience

Received 8 August 2023; Accepted 8 August 2023; Published 9 August 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] S. AlJubayrin, M. Sarfraz, S. A. Ghauri, M. R. Amirzada, and T. Mezgebo Kebedew, "Artificial Bee Colony Based Gabor Parameters Optimizer (ABC-GPO) for Modulation Classification," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 9464633, 9 pages, 2022.

## Research Article

# Artificial Bee Colony Based Gabor Parameters Optimizer (ABC-GPO) for Modulation Classification

Saad AlJubayrin,<sup>1</sup> Mubashar Sarfraz ,<sup>2</sup> Sajjad A. Ghauri ,<sup>3</sup>  
Muhammad Rizwan Amirzada ,<sup>2</sup> and Teweldebrhan Mezgebo Kebedew <sup>4</sup>

<sup>1</sup>Department of Computer Science, College of Computing and Information Technology, Shaqra University, Shaqra, Saudi Arabia

<sup>2</sup>Faculty of Engineering and Computer Science, National University of Modern Languages, Islamabad, Pakistan

<sup>3</sup>School of Engineering & Applied Sciences, ISRA University, Islamabad, Pakistan

<sup>4</sup>Ethio Telecom, Addis Ababa, Ethiopia

Correspondence should be addressed to Teweldebrhan Mezgebo Kebedew; [teweldeberhan.m@ethiotelecom.et](mailto:teweldeberhan.m@ethiotelecom.et)

Received 2 August 2022; Revised 22 August 2022; Accepted 25 August 2022; Published 30 September 2022

Academic Editor: Ateeq Ur Rehman

Copyright © 2022 Saad AlJubayrin et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Modulation classification is one of the essential requirements in the various cognitive radio applications where prior information about the incoming signal is unknown. The modulation classification using a pattern recognition approach can be achieved in 2 modules: first, parameters are extracted from the noisy signal, and then feature selection is carried out using a Gabor filter network (GFN). In the second module, features are exploited for classification purposes. The modulation formats considered for the purpose of classification are BPSK, QPSK, 8PSK, 16PSK, 64PSK, 4FSK, 8FSK, 16FSK, QAM, 8QAM, 16QAM, 32QAM, and 64QAM. The Gabor filter parameters and weights of the adaptive filter are attuned using the Delta rule and recursive least square (RLS) algorithm until the cost function is minimized. In the end, the artificial bee colony (ABC) algorithm is used to optimize the Gabor parameters as well as the classifier's performance. The simulation results show the supremacy of the proposed classifier structure.

## 1. Introduction

Automatic digital modulation classification (ADMC) is to classify the modulation format of the received signal, which has undergone channel effects and noise. For commercial and military communication systems, classification and identification of the modulation format are a significant phase formerly the demodulation at the receiver side. Rapid growth in the commercial wireless communication system demands adaptive, efficient spectrum access algorithms. Software-defined radio (SDR) and later, cognitive radio (CR) are examples of civilian adaptive structures [1–4].

Due to a lack of prior knowledge of the modulation format, ADCM becomes more complex and challenging. The modulation format includes modulation type, symbol duration, frequency deviation, carrier frequency/phase offsets, noise variance, channel amplitude, and so on. In

Figure 1, a typical block diagram of ADCM is shown. The input symbols are first modulated and passed through a channel that adds the additive white Gaussian noise. At the receiver end, the first received signal is preprocessed; demodulation of the received signal and detection of the transmitted signal (information-bearing symbols) are executed after the modulation format is classified [5–7].

*1.1. Contribution of the Article.* The literature review shows that the choice of efficient features and classifier needs to be addressed for performance improvement in the classifier structure. Under the effect of fading channels and AWGN, GFN features for classification of modulation formats BPSK, QPSK, 8PSK, 16PSK, 64PSK, 4FSK, 8FSK, 16FSK, QAM, 8QAM, 16QAM, 32QAM, and 64QAM are presented in this paper. The classifier performance is further optimized by

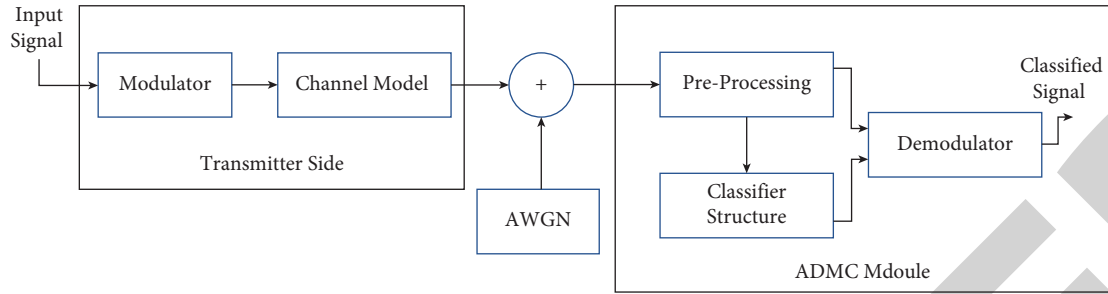


FIGURE 1: Typical block diagram of ADMC.

using the artificial bee colony algorithm (ABC). The performance of the proposed algorithm is evaluated with and without optimization.

**1.2. Organization of the Article.** The rest of the paper is as follows: section 2 contains a comprehensive review of the literature on automatic digital modulation classification. A system model with a proposed classifier structure is explained in section 3. The Gabor filter structure for digital modulation classification and the training and testing of the proposed algorithm are also presented in section 4. The detailed simulation results are carried out in section 4, which shows the supremacy of the proposed classifier. section 5 concludes the paper.

## 2. Related Work

The ADMC can be generally classified into two categories: the likelihood ratio-based decision-theoretic approach and the feature extraction-based pattern recognition approach [8]. In a decision-theoretic approach, the decision is made based on the likelihood function (LF) of the received signal. Once the likelihood function is constituted, the latter classifies the modulation format of the received signal. The process of ADMC in the decision-theoretic approach may be viewed as multiple hypothesis tests or a sequence of pairwise multiple hypothesis tests. To compute the unconditional likelihood, the following are the well-known algorithms in the literature: average likelihood ratio test (ALRT), generalized likelihood ratio test (GLRT), hybrid likelihood ratio test (HLRT), and quasi-hybrid likelihood ratio test (QHLRT) [9–12].

Some related work for the decision-theoretic approach is presented in [13], where the authors use it to determine the modulation format for software-defined radio. A lookup table (LUT)-based classifier is proposed. In [14], only amplitude modulation formats are considered, and the proposed classifier is based on a hybrid maximum likelihood approach. In [15], likelihood algorithms (HLRT, QHLRT) are explored for digital modulation classification. The complexity of HLRT is evaluated, and whether QHLRT provides a reasonable solution is also discussed. The Cramer-Rao upper bound for BPSK and QPSK modulation formats is also employed. In [16], the authors survey the existing techniques for the modulation classification problem.

Feature extraction-based pattern recognition approaches are suboptimal. The FB approach is carried out in two

modules, feature extraction and classifier structure [17]. The features extracted are spectral features, statistical features, cyclo-stationary features, and time-frequency features. They are found in the literature for the FB approach. The features extracted from the PSO-SC with the best clustering radius are shown in [18]. The spectral features are used to classify the 9 analogue and digital modulation formats, and the back propagation neural network is used as a classifier in [19, 20]. The proposed algorithm in [21] uses genetic programming with K-nearest neighbour (GP-KNN) and higher-order cumulants as features to classify the four modulation formats.

A fuzzy logic-based modulation classification is proposed in [22]. The author develops a nonsingleton fuzzy logic classifier by using a fuzzy logic system (FLS). In [23], the author proposed cyclo-stationary-based feature detection for the problem of modulation classification for cognitive radios. The author employed a neural network (NN) and hidden Markov model (HMM)-based classifiers. The spectral features are developed for the classification of ASK, PSK, and FSK using a maximum likelihood decision-based criterion as a classifier [24]. In [25], the author utilises the higher-order cumulants (HOC) as features, and the classifier is based on a support vector machine (SVM). The binary SVM and multiclass SVM are used in conjunction with genetic algorithms, and classifier performance is evaluated with and without optimization.

The spectral features and HOC are extracted, and two multilayer perceptron recognizers, namely a back propagation neural network (BPNN) and resilient back-propagation (RPROP), are used in [26]. The bee's algorithm for optimization of the performance of the classifier is utilised in [27]. The features used are HOC and instantaneous characteristics of digital modulations. The hierarchical SVM is used as a classifier. Under the multipath fading environment, the normalised fourth-order cumulants are used to classify the BPSK, QPSK, and QAM. The Cramer-Rao lower bound is consequent to the features in [28]. The extracted time frequency is used as input to an MLP-based NN for the classification of digital modulation formats in [29, 30].

## 3. Proposed Classifier Structure

**3.1. System Model.** The generalized expression for the received signal is given as follows:

TABLE 1: Symbols and description.

Symbol	Description
$L$	Order of multipath channel
$g_i(t)$	Gabor atom
$\sigma_i$	$i^{\text{th}}$ Gabor scale parameter
$c_i$	$i^{\text{th}}$ Gabor shift parameter
$f_i$	$i^{\text{th}}$ Gabor modulation parameter
$k(n)$	Gain vector
$K(n)$	Weight error correlation
$\Phi(n)$	Time average correlation matrix
$\lambda(n)$	Positive constant

$$r(n) = \sum_{l=0}^L h[n, l]x[n-l] + w(n), \quad (1)$$

where,  $L$  is the order of Multipath channel,  $x(n)$  is the modulated signal,  $w(n)$  is the additive white Gaussian noise, and  $h[n, l]$  is the response of the channel. In a matrix form, the following equation is given as follows:

$$\mathbf{r}_n = \mathbf{H}_n \mathbf{x}_n + \mathbf{w}_n. \quad (2)$$

Table 1 shows the symbols and their descriptions, which are used in the article. The system model of proposed classifier structure, which is an artificial bee colony assisted Gabor filter is shown in Figure 2. The proposed classifier structure classifies the BPSK, QPSK, 8PSK, 16PSK, 64PSK, 4FSK, 8FSK, 16FSK, QAM, 8QAM, 16QAM, 32QAM, and 64QAM efficiently.

**3.2. Gabor Filter Structure.** The Gabor filter-based architecture is a tool for efficient feature extraction from the received signal. The Gabor atom is defined as follows in [31]:

$$g_i(t) = \frac{1}{\sqrt{\sigma_i}} e^{-\pi(t-c_i/\sigma_i)^2} \cos(f_i t), \quad (3)$$

where,  $c$ ,  $\sigma$  and  $f$  are the shift, scale, and modulation parameters, respectively. The output of the  $i^{\text{th}}$  Gabor atom node is corresponding to the received signal is given as follows:

$$\phi_i = \left| r_i \underbrace{\frac{1}{\sqrt{\sigma_i}} e^{-\pi(t-c_i/\sigma_i)^2} \cos(f_i t)}_{g_i(t)} \right|. \quad (4)$$

The GFN output  $\phi_i$  in the input layer is weighted by  $w_i$  i.e.

$$y(n) = \sum_{i=1}^M \phi_i w_i. \quad (5)$$

The difference between the desired response  $d(n)$  and the output  $y(n)$  is the error function denoted by the following equation:

$$e(n) = d(n) - y(n). \quad (6)$$

The extracted features  $c$ ,  $\sigma$  and  $f$  are the Gabor atom parameters and weights of the adaptive filter  $w$  are adjusted until the cost function is minimized. The cost function is the sum of squared errors, which is [31] given as follows:

$$J(n) = [e(n)]^2. \quad (7)$$

**3.3. Training of the Classifier.** Figure 3 shows the training process for ADCM. The two adaptive algorithms are executed by the Gabor filter network in the training phase of ADCM; features are obtained by adjusting the Gabor atom parameters ( $c$ ,  $\sigma$  and  $f$ ) for each modulation format in the first algorithm, while in the second algorithm the weights are adjusted to minimize the error function. The delta rule, which is used to update the parameters of GFN, is taken from [31]. The updating of shift, scale, and frequency parameters are shown in equations (8)–(10).

$$c_i(n+1) = c_i(n) + \eta_c (d(n) - y(n)) w_i \times \left[ \frac{x_i}{\sqrt{\sigma_i}} \cos(f_i t) 2\pi(t - c_i) e^{-\pi(t - c_i/\sigma_i)^2} \right], \quad (8)$$

$$\sigma_i(n+1) = \sigma_i(n) + \eta_\sigma (d(n) - y(n)) w_i \times \left[ \frac{x_i \cos(f_i t)}{\sqrt{\sigma_i}} e^{-\pi(t - c_i/\sigma_i)^2} \left[ \frac{2\pi(t - c_i)^2}{\sigma^2} - \frac{1}{2\sigma_i} \right] \right], \quad (9)$$

$$f_i(n+1) = f_i(n) + \eta_f (d(n) - y(n)) w_i \times \left[ \frac{t}{\sqrt{\sigma_i}} x_i e^{-\pi(t - c_i/\sigma_i)^2} \sin(f_i t) \right]. \quad (10)$$

The weights of the GFN are updated using an RLS filter, and the weight updating equation is shown in equation (13)

$$\mathbf{k}(n) = \frac{\mathbf{k}(n-1)\Phi(n)}{\lambda + \Phi^T(n)\mathbf{K}(n-1)\Phi(n)}, \quad (11)$$

$$\mathbf{K}(n) = \lambda^{-1}\mathbf{K}(n-1) - \lambda^{-1}\mathbf{k}(n)\Phi^T(n)\mathbf{K}(n-1),$$

$$\mathbf{w}(n) = \mathbf{w}(n-1) + \mathbf{k}(n)e(n). \quad (12)$$

**3.4. Testing of the Classifier.** In the testing phase of the proposed classifier structure, the received signal is first serially converted to parallel and then fed to the trained Gabor filter bank. The error has been calculated for each Gabor filter as shown in Figure 4. The minimum error corresponds to the desired modulation format [31].

**3.5. Artificial Bee Colony Optimizer.** The artificial bee colony (ABC) algorithm is used to optimize the Gabor filter parameters ( $c_i$ ,  $\sigma_i$ ,  $f_i$ ) as well as the weights  $w(n)$ . The Gabor parameters and weights are optimized by minimizing the cost function defined in (7). The ( $c$ ,  $\sigma$ ,  $f$ ,  $w$ ) are randomly initialized at the start of the algorithm, and fitness is

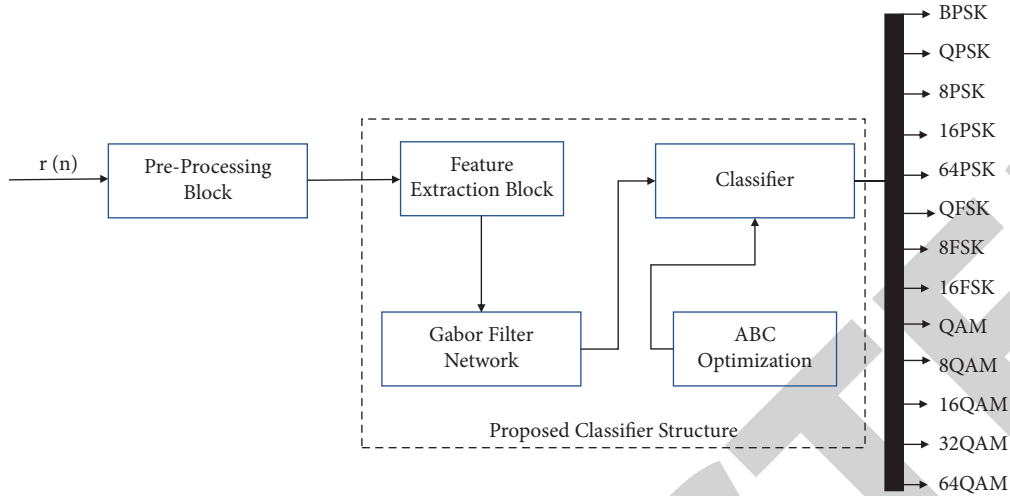


FIGURE 2: System model.

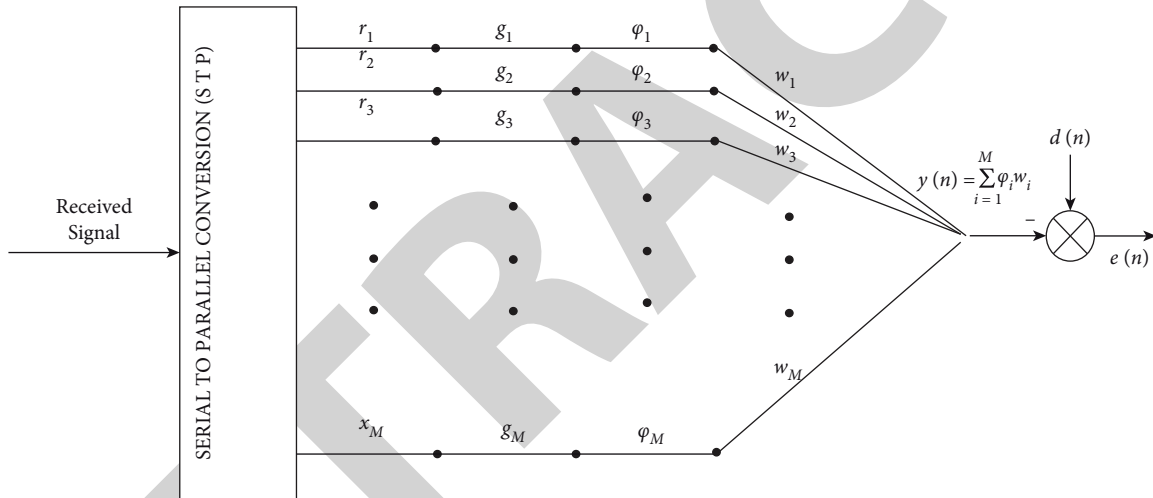


FIGURE 3: Training structure of the Gabor filter for ADCM.

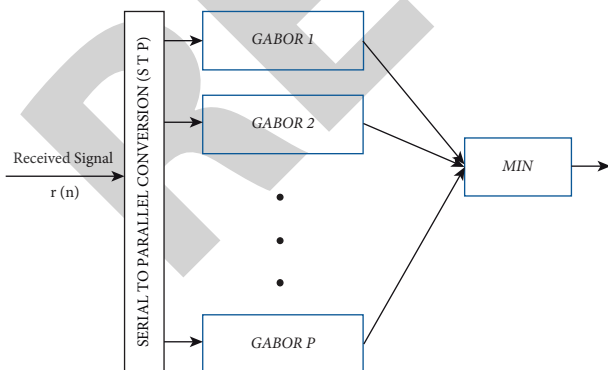


FIGURE 4: Testing of proposed classifier.

evaluated. The ABC optimization adapts to the natural behaviour process of the honeybees. The solutions are updated by searching the neighbouring areas through three different processes that are carried out by employer bees,

onlooker bees, and scout bees. Algorithm 1 presents the brief steps for the ABC-GPO algorithm.

#### 4. Experimental Classification Results and Analysis

The performance of the proposed classifier using the optimized Gabor filter features is evaluated in this section. The thirteen modulation formats are considered for classification, i.e., BPSK, QPSK, 8PSK, 16PSK, 64PSK, 4FSK, 8FSK, 16FSK, QAM, 8QAM, 16QAM, 32QAM, and 64 QAM. For simplicity, these modulated signals are replaced by P1, P2, P3, P4, P5, P6, P7, P8, P9, P10, P11, P12, and P13. All the simulations are done in MATLAB R2020a. The performance metric is the percentage classification accuracy (PCA). A total of 100000 realizations have been taken with 1024 no. of samples. The data set has been divided into training, validation, and testing samples, i.e., 80%, 10%, and 10%.

```

(1) Initialize
(2) Gabor parameters:  $c, \sigma, f, w$ 
(3) while iterations == true do
(4)   evaluate fitness through (7)
(5)   apply employer bees phase
(6)   apply onlooker bees phase
(7)   apply scout bees phase
(8)   update Gabor parameters
(9)   if eq. ((7)  $\rightarrow$  0 (cost function is minimized) then,
(10)     break;
(11)   end
(12) end
(13) return  $c, \sigma, f, w$ 

```

ALGORITHM 1: ABC-based Gabor parameters optimizer (ABC-GPO).

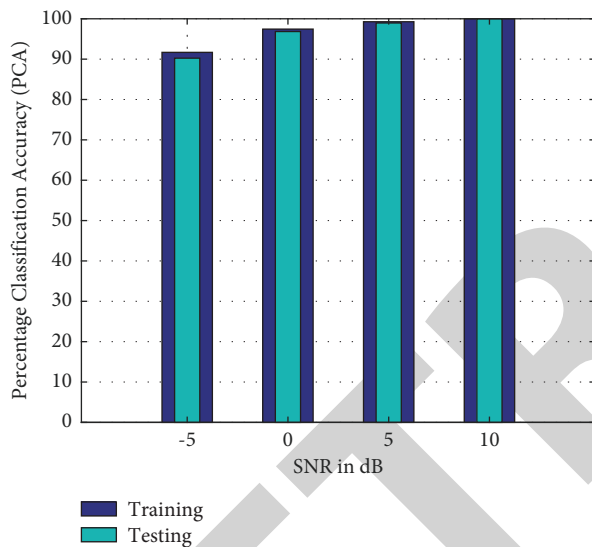


FIGURE 5: Training and testing of the proposed classifier on the AWGN model.

**4.1. PCA on AWGN Channel.** Figure 5 shows the training and testing performance of the classifier on the AWGN channel for all considered modulation formats. The training of the proposed GFN is carried out at different SNRs. The percentage classification accuracy for the training and testing of the classifier is approximately 90% at an SNR of  $-5$  dB. At 10 dB SNR, the classification accuracy approaches 100% for the training, and testing of a classifier. The total number of samples taken is 1024, with 10,000 realizations.

**4.2. PCA on Rician Channel.** Figure 6 shows the training and testing performance of the classifier on the Rician fading channel with AWGN. The percentage classification accuracy for the training and testing of the classifier is approximately 99.42% and 98.50% at an SNR of 10 dB, respectively, which is relatively low as compared to the AWGN channel, while PCA is below 90% at  $-5$  dB of SNR.

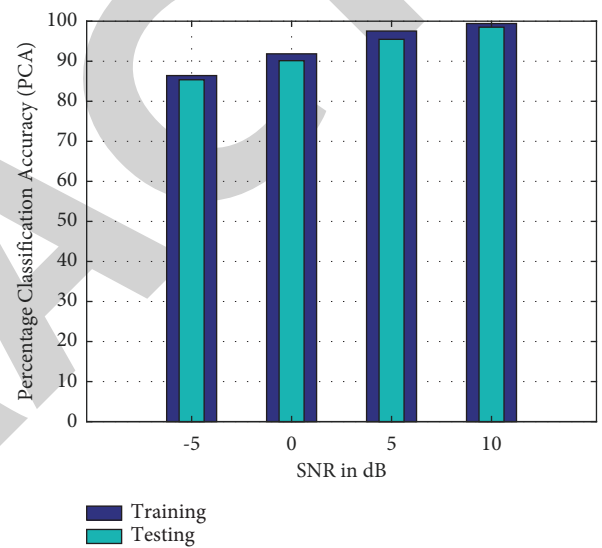


FIGURE 6: Training and testing of the proposed classifier on the Rician model.

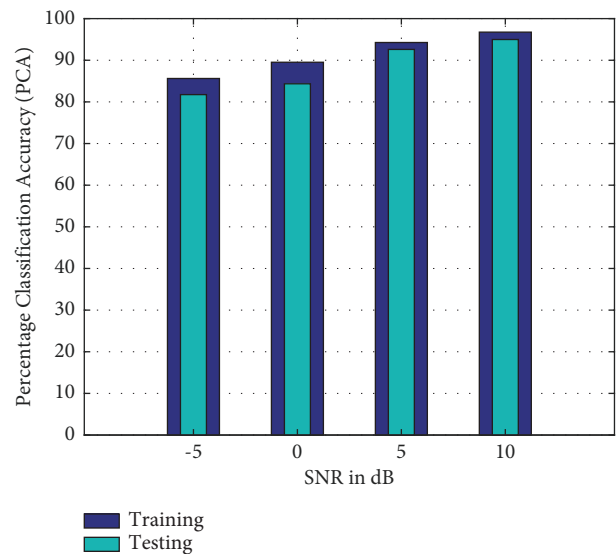


FIGURE 7: Training and testing of the proposed classifier on the Rayleigh model.



TABLE 3: Percentage classification performance of ADMC with optimization at a SNR of 8 dB.

	$P_1$	$P_2$	$P_3$	$P_4$	$P_5$	$P_6$	$P_7$	$P_8$	$P_9$	$P_{10}$	$P_{11}$	$P_{12}$	$P_{13}$
$P_1$	100												
$P_1$		99											
$P_1$			98.5										
$P_1$				98.8									
$P_1$					99.2								
$P_1$						99.1							
$P_1$							99.6						
$P_1$								99.4					
$P_1$									98.4				
$P_1$										99.1			
$P_1$											99.4		
$P_1$												98.4	
$P_1$													98.8

TABLE 4: Comparison with the existing techniques.

Method and reference	Number of modulation	Classification accuracy in % at SNR = 10 dB	No. of features
Pdf of the received signal, zero crossing [32]	PSK, FSK (7)	98 (15 dB of SNR)	4
Spectral features [33]	PSK, FSK (6)	90	5
Spectral features and moments [34]	PSK, FSK (9)	96 (15 dB of SNR)	9
Constellation shape [35]	PSK, QAM (3)	90 (5 dB of SNR)	4
HOC [36]	PSK, QAM (8)	96	2
Neural networks [37]	PSK, FSK, QAM (7)	93 (8 dB of SNR)	5
Back propagation (BP) algo., resilient BP algo [26]	PSK, FSK, QAM (10)	89.96, 99.95	17
Time frequency features [29]	FSK, PSK (6)	97.64	4
Normalized cummulants [28]	PSK, QAM (5)	97.5	1
Cummulants features, multiclass SVM [25]	PSK, QAM (9)	97.56	12
Spectral features and moments [30]	FSK, QAM (12)	99.95	8
HOC and spectral features, multiclass SVM [27]	PSK, QAM, FSK (11)	97.45	11
Gabor filter based features [31]	PSK, FSK, QAM (18)	100	3
HOC and SVM [38]	QAM (6)	98.7 (5 dB of SNR)	11
HOC and polynomial classifier [39]	PSK (5)	98	10
Proposed optimized classifier	PSK, FSK, QAM (13)	100 (8 dB of SNR)	3

without optimization is the complexity of the GFN, which is given in [40] as follows:

$$O_{GFN} = O(N \log N), \quad (13)$$

where  $N$  is the number of samples, however, the complexity of the proposed algorithm with optimization is as follows:

$$O_{ABC-GFN} = O(\text{Iterations} * \text{No. of Bees} * \text{eq. (7)}). \quad (14)$$

## 5. Conclusion

The automatic digital modulation classification using GFN is considered in this research article. The extracted features are GF features for the considered BPSK, QPSK, 8PSK, 16PSK, 64PSK, 4FSK, 8FSK, 16FSK, QAM, 8QAM, 16QAM, 32QAM, and 64QAM modulation formats. The bee colony algorithm is used to optimize the algorithm's results. The simulation results show the performance of the classifier with and without optimization of the algorithm. The performance fading channels is compared to the state of the art of existing techniques.

## Data Availability

This article does not meet the criteria for data sharing since no datasets were generated or analyzed.

## Conflicts of Interest

The authors declare that there are no conflicts of interest.

## References

- [1] S. Alam, M. F. Sohail, S. A. Ghauri, I. Qureshi, and N. Aqdas, "Cognitive radio based smart grid communication network," *Renewable and Sustainable Energy Reviews*, vol. 72, pp. 535–548, 2017.
- [2] A. U. Rehman, R. A. Naqvi, A. Rehman, A. Paul, M. T. Sadiq, and D. Hussain, "A trustworthy snot aware mechanism as an enabler for citizen services in smart cities," *Electronics*, vol. 9, no. 6, p. 918, 2020.
- [3] M. K. Ehsan, "Performance analysis of the probabilistic models of ism data traffic in cognitive radio enabled radio environments," *IEEE Access*, vol. 8, pp. 140–150, 2020.



- [4] M. K. Ehsan and D. Dahlhaus, "Statistical modeling of ism data traffic in indoor environments for cognitive radio systems," in *Proceedings of the 2015 3rd International Conference on Digital Information, Networking, and Wireless Communications (DINWC)*, pp. 88–93, IEEE, Moscow, Russia, July 2015.
- [5] M. Sarfraz, S. Alam, S. A. Ghauri et al., "Random graph-based m-qam classification for mimo systems," *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 9419764, 10 pages, 2022.
- [6] R. Khan, Q. Yang, I. Ullah et al., "3d convolutional neural networks based automatic modulation classification in the presence of channel noise," *IET Communications*, vol. 16, no. 5, pp. 497–509, 2022.
- [7] S. A. Ghauri and I. Mansoor Qureshi, "M-PAM signals classification using modified gabor filter network," *Mathematical Problems in Engineering*, vol. 2015, 10 pages, 2015.
- [8] S. I. H. Shah, S. Alam, S. A. Ghauri, A. Hussain, and F. Ahmed Ansari, "A novel hybrid cuckoo search-extreme learning machine approach for modulation classification," *IEEE Access*, vol. 7, pp. 90 525–590 537, 2019.
- [9] J. L. Xu, W. Su, and M. Zhou, "Likelihood-ratio approaches to automatic modulation classification," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 41, no. 4, pp. 455–469, 2011.
- [10] C.-Y. Huan and A. Polydoros, "Likelihood methods for mpsk modulation classification," *IEEE Transactions on Communications*, vol. 43, no. 2/3/4, pp. 1493–1504, 1995.
- [11] O. A. Dobre and F. Hameed, "Likelihood-based algorithms for linear digital modulation classification in fading channels," in *Proceedings of the 2006 Canadian conference on electrical and computer engineering*, pp. 1347–1350, IEEE, Ottawa, Canada, May 2006.
- [12] M. Derakhtian, A. Tadaion, and S. Gazor, "Modulation classification of linearly modulated signals in slow flat fading channels," *IET Signal Processing*, vol. 5, no. 5, pp. 443–450, 2011.
- [13] W. Su, J. L. Xu, and M. Zhou, "Real-time modulation classification based on maximum likelihood," *IEEE Communications Letters*, vol. 12, no. 11, pp. 801–803, 2008.
- [14] O. Ozdemir, R. Li, and P. K. Varshney, "Hybrid maximum likelihood modulation classification using multiple radios," *IEEE Communications Letters*, vol. 17, no. 10, pp. 1889–1892, 2013.
- [15] F. Hameed, O. A. Dobre, and D. C. Popescu, "On the likelihood-based approach to modulation classification," *IEEE Transactions on Wireless Communications*, vol. 8, no. 12, pp. 5884–5892, 2009.
- [16] O. A. Dobre, A. Abdi, Y. Bar-Ness, and W. Su, "Survey of automatic modulation classification techniques: classical approaches and new trends," *IET Communications*, vol. 1, no. 2, pp. 137–156, 2007.
- [17] S. A. Ghauri and I. Qureshi, "Automatic classification of digital modulated signals using linear discriminant analysis on awgn channel," in *Proceedings of the 1st International Conference on Information and Communication Technology Trends (ICICTT'13)*, Yogyakarta, Indonesia, October 2013.
- [18] L. Yan-ling, L. Bing-bing, and Y. Chang-yi, "Modulation classification of mqam signals using particle swarm optimization and subtractive clustering," in *Proceedings of the IEEE 10th International Conference on Signal Processing Proceedings*, pp. 1537–1540, IEEE, Piraeus, Greece, July 2010.
- [19] J. J. Popoola and R. van Olst, "Automatic classification of combined analog and digital modulation schemes using feedforward neural network," in *Proceedings of the IEEE African'11*, pp. 1–6, IEEE, Victoria Falls, Zambia, September 2011.
- [20] V. D. Orlic and M. L. Dukic, "Automatic modulation classification: sixth-order cumulant features as a solution for real-world challenges," in *Proceedings of the 2012 20th Telecommunications Forum (TELFOR)*, pp. 392–399, IEEE, Belgrade, Serbia, November 2012.
- [21] M. W. Aslam, Z. Zhu, and A. K. Nandi, "Automatic digital modulation classification using genetic programming with k-nearest neighbor," in *Proceedings of the 2010-Milcom 2010 Military Communications Conference*, pp. 1731–1736, IEEE, San Jose, California, USA, November 2010.
- [22] W. Wei and J. M. Mendel, "A fuzzy logic method for modulation classification in nonideal environments," *IEEE Transactions on Fuzzy Systems*, vol. 7, no. 3, pp. 333–344, 1999.
- [23] B. Ramkumar, "Automatic modulation classification for cognitive radios using cyclic feature detection," *IEEE Circuits and Systems Magazine*, vol. 9, no. 2, pp. 27–45, 2009.
- [24] E. Azzouz and A. Nandi, "Procedure for automatic recognition of analogue and digital modulations," *IEEE Proceedings - Communications*, vol. 143, no. 5, pp. 259–266, 1996.
- [25] A. Ebrhimzadeh, H. Azimi, and H. M. Naeemi, "Classification of communication signals using an optimized classifier and efficient features," *Arabian Journal for Science and Engineering*, vol. 35, 2010.
- [26] M. D. Wong and A. K. Nandi, "Automatic digital modulation recognition using artificial neural network and genetic algorithm," *Signal Processing*, vol. 84, no. 2, pp. 351–365, 2004.
- [27] A. E. Sherme, "A novel method for automatic modulation recognition," *Applied Soft Computing*, vol. 12, no. 1, pp. 453–461, 2012.
- [28] H.-C. Wu, M. Saquib, and Z. Yun, "Novel automatic modulation classification using cumulant features for communications via multipath channels," *IEEE Transactions on Wireless Communications*, vol. 7, no. 8, pp. 3098–3105, 2008.
- [29] Y. Ye and M. Wenbo, "Digital modulation classification using multi-layer perceptron and time-frequency features," *Journal of Systems Engineering and Electronics*, vol. 18, no. 2, pp. 249–254, 2007.
- [30] J. J. Popoola and R. Olst, "A novel modulation-sensing method," *IEEE Vehicular Technology Magazine*, vol. 6, no. 3, pp. 60–69, 2011.
- [31] S. A. Ghauri, I. M. Qureshi, T. A. Cheema, and A. N. Malik, "A novel modulation classification approach using gabor filter network," *The Scientific World Journal*, vol. 2014, 14 pages, 2014.
- [32] S.-Z. Hsue and S. S. Soliman, "Automatic modulation classification using zero crossing," *IEEE Proceedings F (Radar and Signal Processing)*, vol. 137, no. 6, pp. 459–464, 1990.
- [33] E. E. Azzouz and A. K. Nandi, "Automatic identification of digital modulation types," *Signal Processing*, vol. 47, no. 1, pp. 55–69, 1995.
- [34] A. K. Nandi and E. E. Azzouz, "Algorithms for automatic modulation recognition of communication signals," *IEEE Transactions on Communications*, vol. 46, no. 4, pp. 431–436, 1998.
- [35] B. G. Mobasser, "Digital modulation classification using constellation shape," *Signal Processing*, vol. 80, no. 2, pp. 251–277, 2000.
- [36] A. Swami and B. M. Sadler, "Hierarchical digital modulation classification using cumulants," *IEEE Transactions on Communications*, vol. 48, no. 3, pp. 416–429, 2000.

## Research Article

# A Novel Framework for Fog-Assisted Smart Healthcare System with Workload Optimization

**Ahmed A. H. Abdellatif** <sup>1,2</sup>, **Aman Singh** <sup>2,3</sup>, **Abdulaziz Aldribi**,<sup>2,4</sup>  
**Arturo Ortega-Mansilla**,<sup>3,5,6</sup> and **Muhammad Ibrahim** <sup>7</sup>

<sup>1</sup>Department of Pharmaceutics, College of Pharmacy, Qassim University Buraydah, Buraydah, Saudi Arabia

<sup>2</sup>Prince Faisal Bin Mishaal Artificial Intelligence Chair, Qassim University, Buraydah, Saudi Arabia

<sup>3</sup>Higher Polytechnic School, Universidad Europea del Atlántico, C/Isabel Torres 21, Santander 39011, Spain

<sup>4</sup>Department of Computer Science, College of Computer, Qassim University, Buraydah, Saudi Arabia

<sup>5</sup>Department of Project Management, Universidad Internacional Iberoamericana, Arecibo 00613, Puerto Rico, USA

<sup>6</sup>Department of Project Management, Universidad Internacional Iberoamericana, Campeche 24560, Mexico

<sup>7</sup>Department of Information Technology, University of Haripur, Haripur 22620, Pakistan

Correspondence should be addressed to Ahmed A. H. Abdellatif; [a.abdellatif@qu.edu.sa](mailto:a.abdellatif@qu.edu.sa) and Muhammad Ibrahim; [ibrahimmayar@uoh.edu.pk](mailto:ibrahimmayar@uoh.edu.pk)

Received 28 June 2022; Revised 22 July 2022; Accepted 4 August 2022; Published 29 September 2022

Academic Editor: Ateeq Ur Rehman

Copyright © 2022 Ahmed A. H. Abdellatif et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Fog-assisted and IoT-enabled smart healthcare system with rapid response rates is the major area of concern now a days. Dynamic and heterogeneous fog networks are difficult to manage and a considerable amount of overhead could be realized while managing ever increasing load on foglets. Fog computing plays a vital role in managing ever increasing processing demands from diverse IoT-based applications. Smart healthcare systems work with the assistance of sensor-based devices and automatic data collection and processing can speed up overall system functionality. In the proposed work, a novel framework for smart health care is presented where a series of activities are performed with prime objective of reducing latency and execution time. Principal component analysis is used for feature reduction and support vector machines with radial basis function kernel is used for classification purpose. Workload optimization on the fog nodes is implemented using genetic algorithm. Data collection process also involves preprocessing as a leading step for generating cleaner data. Amalgamation of intelligent and optimization techniques in the presented framework certainly improves the efficiency of the overall system. Experimental results reveal that proposed work outperforms the existing fog-assisted smart healthcare systems in terms of latency, execution time, overall system accuracy, and system stability.

## 1. Introduction

With the emergence of IoT devices and related technology, smart healthcare systems have evolved tremendously. The existence of these systems and their associated benefits have also paced up the processing speed, and the accuracy levels of the results could be seen on a higher side. Fog computing has also emerged as a supporting technology in this dynamic era to fulfill the current requirements. It serves as an intermediate between the

cloud and the IoT layer [1]. Intelligent healthcare systems are latency-sensitive, and sending all the requests to the cloud layer can cause significant delays. The fog layer solves the above-stated problem, and non-critical tasks could be served immediately so that only critical tasks could be sent to the cloud layer. Cloud layer can provide resources in large amounts, which could be required to process critical tasks. Processing of non-critical tasks at the fog layer can reduce a considerable amount of overhead on the cloud layer [2].

Fog computing paradigm works by processing the tasks generated at the IoT layer. As, this layer works as the middle layer, it can produce results quickly which could be used by another set of operations. Fog servers possess less resources, due to which it becomes difficult to manage large set of tasks. For resolving this issue, proper task management should be applied with suitable optimization procedures. Fog computing offers its assistance in a diversified set of applications like: Smart City, Smart Grid, Smart Industry, Smart healthcare etc., With the increasing attention towards healthcare sector, fog computing is playing a pivotal role and also is proving its effectiveness in this area. In the last couple of years, enormous amount of research has been laid on smart healthcare with promising results.

Fog computing enables location awareness, support for user accessibility, real-time interactions, low latency, high scalability, and interoperability, which cloud computing systems could not support. Because most of the information is being processed locally in the fog nodes, only summarized data will be transported to the cloud, resulting in significant bandwidth savings. It will decrease latency, delays, and packet loss [3]. Fog computing requires effective resource management. The three-tier mechanism of fog computing is presented in Figure 1.

Managing workload in the fog computing paradigm is a major challenge, as numerous requests are generated from the IoT layer at the same point of time. Timely management of such requests becomes crucial because of the limited number of resources at the fog layer. In addition to it, some concern must be given to critical smart healthcare systems as well, where delay could not be tolerated because of lack of load management. This issue could further put a negative impact on the related performance parameters such as: response time, energy consumption, cost. Proper load balancing works to accomplish multiple objectives such as: optimal throughput and minimized response time which further improves latency and energy consumption [4].

*1.1. Limitations and Research Goals.* This research work is the extension to the work that has been done over the years in the field of smart healthcare systems. Various limitations from the existing systems are extracted, and they all can be summarized as following [1, 2, 4]:

- (i) lack of data preprocessing—large amount of sensor data is generated which is required to be pre-processed before actually going for the final prediction and analysis.
- (ii) Heavy workload on fog servers—as sensors can generate various tasks and huge amount of data is required to be processed, fog servers could be easily overloaded, and it can hamper their performance and latency could be achieved.
- (iii) Heavy energy consumption—due to overloaded fog servers, there could be a chance of heavy energy consumption which can reduce the overall system efficiency.
- (iv) High Latency—improper data processing and resource management can cause considerable amount of delays, therefore high latency could be observed.
- (v) Dynamic requests—smart environments are highly dynamic in nature and sometimes it becomes tedious task to manage the ever-changing requests, as fog servers are mostly configured for the static set of requests from the IoT layer.
- (vi) Unnecessary features for task classification—there are various features or data attributes which are unnecessary and they play a minimal role in the final classification of the task category.

The above section is signifying the need of extending the existing research work to formulate new procedures which can minimize the effects of identified limitations. A summarized view of various objectives with respect to desired architecture for smart healthcare system is presented below. The first objective is to perform data preprocessing to achieve cleaner and accurate data. The second objective is to minimize workload on the overloaded fog servers which can lead to fulfillment of the third objective of reducing latency and lastly it can handle dynamic requests from the IoT layer. The abovementioned objectives are achieved using the following phases of designed architecture, such as:

- (a) implementation of data preprocessing techniques for noise removal and data filtering.
- (b) Application of dimensionality reduction technique: principal component analysis (PCA) for reducing unwanted features.
- (c) Classification of tasks as critical and non-critical using SVM (support vector machine) for performing the processing at cloud and fog layers.
- (d) Managing load allocation on the fog servers using nature-inspired optimization algorithm, i.e., genetic algorithm

These phases work in a repetitive manner to achieve the desired goals and to provide solutions to the above-mentioned problems. The rest of the paper is organized like, Section 2 is focused on the existing work with respect to smart healthcare systems. Section 3 is focused on the proposed work and the methodology to achieve the desired research goals. Experimental settings and results are discussed in Section 4, and conclusion part sums up the entire work in Section 5.

## 2. Related Work

Over the years, noticeable research is performed to streamline various operations of smart healthcare systems. Multiple approaches are designed with fog and cloud layers to reduce the service delay and enhance the processing capabilities. For instance, framework using ensemble deep learning is developed for automatic heart disease analysis. Performance of the proposed system is evaluated using various parameters, and simulation is performed with the help of FogBus. The proposed work was tested on the real-

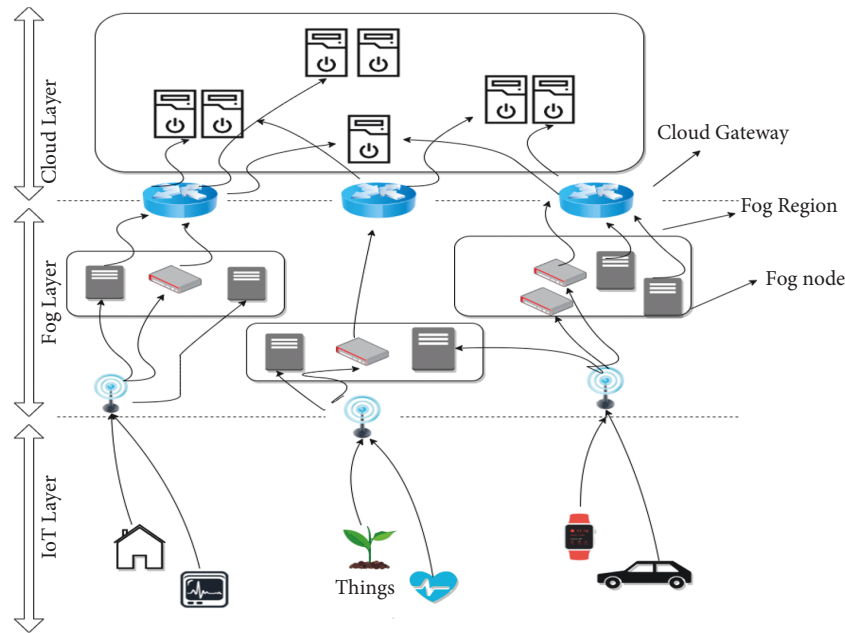


FIGURE 1: Three-tier mechanism of fog computing [1].

time heart patient dataset, and good accuracy levels were observed. It considers various parameters for validation purpose and therefore successfully overpowers the existing systems [5]. A personalized healthcare model was also designed with the help of dynamic programming approach for providing quick service as per patient's health status. The overall model is known as smart treatment for personalized healthcare which optimizes the healthcare services through intelligent agents [6]. Some hybrid approaches were also designed as a solution to the problems incurred in the cloud-only systems. Features of naïve Bayes and firefly algorithms were combined to perform the health prediction of elderly patients [7].

Gait-based system for older adults was presented where acceleration-based gait recognition method was implemented for increasing the recognition performance of IoT-based devices. It uses PCA (principal component analysis) technique for dimensionality reduction which further helps in the identity recognition. The mentioned IoT devices are the wearable devices for the patients which further sends the patient's data to the cloud layer for processing [8]. Real-time health-monitoring system was also presented based on the four-level remote triage and package localization. A set of 12 hospitals were selected for analysis, and later their ranking was performed. The major parameter for ranking was based on the services provided with respect to healthcare. Various types of sensors were deployed to receive patient's data, and the dataset was formulated with a variety of symptoms [9]. Reversed naïve Bayesian logic was also implemented to enhance the classification accuracy of biomedical data. The major aim was to ignore those data instances which can reduce the classification accuracy. So, after reducing the irrelevant data instances, identified noise instances were used for training the model. This approach can reduce the overall training effort to a large extent, and biomedical data classification could be achieved. Various datasets from UCI

were taken for the validation purpose and promising results were seen [10].

Cloud-based health care system was designed which used the concept of digital twin. This approach helps in performing medical simulations for achieving faster and accurate results. It works in three parts, such as physical and virtual objects followed by healthcare data. All medical devices are the part of physical object, whereas virtual object focuses on modelling of medical devices. Healthcare data is the outcome from the sensors and wearable devices. This technique can decrease efforts required for real-time supervision, and the overall system accuracy could be enhanced [11]. Medical data monitoring system was also presented for sensor fault detection. The main objective of this system was to identify those sensor nodes which can generate false results. So, for achieving this, a Bayesian network model was designed based on conditional probability distribution. After implementing the proposed idea, fault detection accuracy increased to a greater extent [12].

Another system which deployed hierarchical computing architecture and convolutional neural network (CNN) was presented which works with an objective of real-time monitoring of patients and performs classification on the ECG data. The main purpose of this work was to enhance the performance accuracy by exploiting the deep learning methods on a real-time dataset. Parameters such as: response time and accuracy were chosen for the validation of the proposed work, and promising results were observed [13]. A novel cloud-based healthcare system was presented with prime purpose to monitor health and perform diagnosis. Cloud layer was presented to perform logic processing using deep-learning-based architecture. CNN-based model was used to achieve desired results. Data preprocessing mechanism was also implemented to assist the working methodology of proposed system [14].

Fog-assisted health monitoring system was presented which used the concept of temporal mining to generate the health index for a patient. It used Bayesian belief network classifier to train the system. Dataset of various users in smart homes was constructed and variety of parameters were involved to test the accuracy of the proposed system. It was also capable of generating alarms in critical situations [15]. A fog-cloud-computing-based architecture was presented, where user health was monitored in smart office environment. It used the severity index to keep track of the health and generated alerts if situation degrades. It focused on accumulating data from various sources like: health, environment, meal, physical postures. Comparison was also performed with various baseline algorithms such as: K-NN, ANN, and SVM, and the proposed work based on probabilistic measures outperformed them [16].

Fog-assisted healthcare system was proposed, which combined medical signals and sensor data to formulate a remote pain monitoring system. This work aimed to reduce latency and network consumption and validate the effectiveness of proposed work after comparing with cloud-based systems. The two types of signals, sEMG and ECG, were processed to compute the pain levels of the patients. Authors have justified the usage of fog-computing-based architecture in context of computations and storage on the edge layer which is closer to the biomedical sensors, hence producing timely results [17]. For managing heterogeneous fog computing environment and inconsistent communication between fog and the edge devices, an efficient approach was presented where application placement could be done on the network nodes. It tried to improve latency and final network consumption [18]. Another approach was designed for mapping sensor nodes to its parent fog node effectively, so that overall system performance could be improved. Authors designed this approach to handle varied computing capabilities of fog nodes and look to reduce delay and improve network utilization [19].

Another healthcare system based on complex event processing was presented where data sources were combined at certain intervals to generate timely responses. The major objective was to obtain a higher response rate as compared to the traditional healthcare systems [20]. Tri-fog health [21] was also presented which used a series of techniques to overpower the existing healthcare systems. It eliminated the faulty data in the wearable layer itself. It also performed redundant data elimination using FaMOORA algorithm. Furthermore, it used SpikeQ-Net for training the system. It was implemented as a hybrid machine learning algorithm. It also performed fog offloading using multi-objective spotted hyena optimization algorithm. As it performed a series of activities in wearable and fog layer, this architecture had provided promising results.

After going through related work, various strengths and limitations were identified and their summary is presented in Table 1.

### 3. Proposed System Framework

In this section, the proposed framework is defined which works in four phases to perform the required set of actions.

*3.1. Problem Statement.* Smart healthcare systems generate numerous amounts of tasks and heavy load which is required to be processed frequently, so that, instant decisions could be taken. Over the years, many computational architectures were proposed with an objective of reducing latency and execution time. But these systems were lacking refinements and optimizations at certain levels. This proposed work was directed to solve stated problem where every layer in the three-tier architecture was intended to do some tasks; hence, overall system was regulated for achieving desired goals. Moreover, heterogeneous fog nodes and varied computational capabilities were difficult to manage, so, constant monitoring and regular optimizations were required which can improvise overall system's efficiency.

*3.2. System Overview.* The proposed system framework is designed to work in three layers, i.e., the IoT layer, the fog layer, and the cloud layer. The IoT layer consists of various types of sensors for collecting the medical data which is preprocessed for removing noise and various outliers. Final preprocessed data is passed on to the set of fog nodes (F1, F2, F3. . .) where dimensionality reduction procedure works and only necessary features are extracted. This step reduces the effort and overall performance of the fog layer is enhanced. The fog layer also implements the task classification technique for categorizing critical and non-critical tasks. Furthermore, optimization method is also incorporated for overloaded fog nodes. For instance, if any fog node is receiving bulk of requests, that could be offloaded and forwarded to the idle foglets.

Once fog layer is done with its all functionalities, the critical tasks are sent to the cloud layer for the final processing and the responses are generated for necessary action. Overall system architecture could be implemented for variety of scenarios, but it will be more suitable for the healthcare systems where even the minor delays are not acceptable and responses should be processed rapidly. Figure 2 is demonstrating the same scenario discussed above. Each layer plays its individual role and moreover series of functions are performed as per the requirement of the proposed architecture. Table 2 highlights the series of steps performed in each phase along with the required technique implemented for the same.

*3.3. System Description.* A brief overview of the proposed system is presented in Figure 2. Every layer in the proposed architecture consists of a series of activities which can play their part in increasing the overall system performance. Table 2 demonstrates the specific set of techniques used in each layer for implementing the proposed methodology.

From Table 2, we can see the list of activities performed in each layer, along with that, the techniques and resource requirement are also mentioned. Every technique or mentioned algorithm is customized as per the requirement of a smart healthcare system. Following subsections are covering the complete details of these activities.

TABLE 1: Strengths and limitations of related work.

Presented work	Strengths	Limitations
HealthFog [5]	Enhancement in Quality of service and prediction accuracy	(i) It could only work on single domain of diseases. (ii) Proposed architecture is designed for one application area only
STPH [6]	Efficient medical services	(i) Higher computational complexity (ii) Lack of medical offering system
HAAL-NBFA [7]	(i) Implementation of safe-fail module (ii) Minimum feature selection	Proposed framework can monitor limited set of diseases
Fog-BBN [15]	Provides assistance for remote patient health monitoring	(i) Event severity level checking is missing (ii) Requirement of real time alert generation
Fog-Smart Office [16]	Severity Index calculation which reflects the impact of various activities in smart office environment	(i) Transmission of varied types of events in a common and adaptable format (ii) Lesser network load efficiency
Fog-CPE [20]	Minimization of time delay	Bi-directional coordination is missing
Tri-Fog [21] Health	(i) Elimination of faulty data (ii) Removal of redundant data (iii) Data processing using various attributes	(i) Proposed system does not work on specific set of diseases (ii) Security aspects are not covered

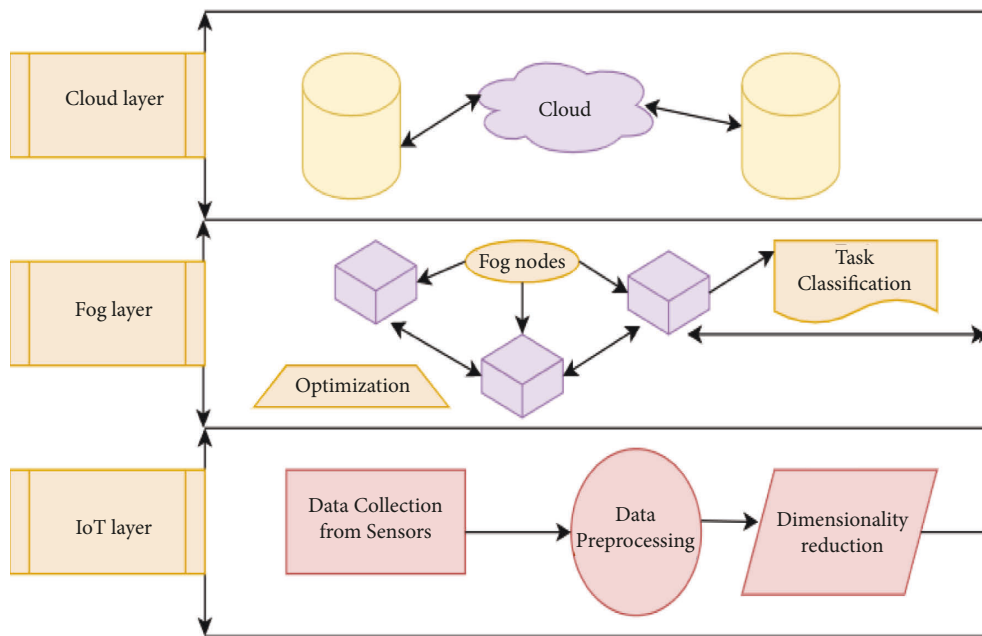


FIGURE 2: Overview of the proposed system.

TABLE 2: Layer-wise functionalities and techniques.

Layer	Functionalities	Techniques/Resources
IoT layer	(i) Data collection	(i) Biomedical sensors
	(ii) Data Preprocessing	(ii) Noise removal and data cleaning
	(iii) Dimensionality reduction	(iii) Principal component analysis
Fog layer	(i) Task classification	(i) Support vector machines
	(ii) Workload optimization	(ii) Genetic algorithm
Cloud layer	(i) Analysis of critical tasks	(i) VMs, PMs

3.4. *Data Collection.* This is the initial step where patient data from various biomedical sensors will be collected and dataset will be formulated. Common parameters, such as blood pressure, temperature, heart rate, Oxygen levels, will be considered, and dataset will be updated after regular intervals. The collected data will be maintained in a file

which will be populated with the parameter values on regular basis. Repository of dataset will be maintained online.

3.5. *Data Preprocessing.* The automated data collection process can gather noisy data which is required to be cleaned

by using some procedures, such as filling up of missing values, identification of outliers, and removing inconsistent values. Routines for performing these all these steps are written and cleaned data is sent for further processing. Sensor data could be easily misunderstood by the receiving algorithm and misclassification could be done. Moreover, various outliers in the data can produce false warnings, and overall system performance will be dropped. Some missing values could also be observed, if due to some reasons, any sensor has stopped working or it is damaged. In these cases, it would be better to have the average values, rather than going for blank set of values. Inconsistent data patterns could also be experienced with faulty sensors, so there should be a mechanism to identify those patterns and look for some countermeasures.

**3.6. Dimensionality Reduction.** This is one of the crucial steps of the methodology wherein only the most important features are extracted and unwanted features are dropped. This step can reduce a considerable amount of effort incurred, and overall system efficiency could be enhanced. PCA (principal component analysis) technique is used to perform this task.

**3.6.1. Principal Component Analysis Model.** It is a multivariate technique whose major role is to analyse the provided data table and to identify the principal components from it. PCA focuses on reducing the number of features without actually reducing the accuracy levels. The major objective is to remove the extraneous variables which can hamper the overall performance. In the proposed system also, this technique is applied to reduce the extraneous medical features, so that rapid classification could be done. It works in a series of steps, such as the first step is to provide standardization for the initial variables. This is done to reduce the variance in the data values. After this step, a covariance matrix is computed to analyse the correlation among various variables. The next step involves the computation of eigen values and vectors to give away the principal components. This step is followed by another step where feature vector is generated. All principal components are not significant, so it becomes necessary to discard some of them. Only significant PCs are taken into consideration to generate the final feature vector. Following equation denotes the final dataset which consists of a reduced number of features decided as per selected eigen vectors [22].

$$FD = FV^P * SOD^P, \quad (1)$$

FD: final dataset, FV: feature vector, SOD: standardized original dataset.

Major tasks of PCA involves the extraction of key information from the dataset and to simplify the complex data description. It can also analyse the patterns of observations and variables associated with them. The extracted information is in the form of principal components which are also known as orthogonal variables. The task of achieving

necessary information/or lower dimensions is achieved using transformation of given data, i.e.,

$$Y = \{y_1, y_2, \dots, y_N\}. \quad (2)$$

from  $S^A$ [Higher dimension] to  $S^B$ [Lower dimension]. In (2), N denotes the maximum number of instances, and  $y_i$  denotes  $i^{\text{th}}$  instance. Further, principal components are derived, and every component has different levels of data variance. For example, the very first component will exhibit maximum data variance, whereas the second one represents next maximum variance and the process of principal components computation continues. There are various methods of calculating principal components such as: covariance-matrix-based and singular-value decomposition. The proposed framework implements covariance-matrix-based methodology to calculate principal components and furthermore dimensionality reduction could be achieved. Covariance matrix (equation (3)) represents various essential things, such as the diagonal elements represent variance among two variables and further if there is a positive element available in the matrix, it reflects the positive correlation and a negative value will highlight the negative correlation. A zero value represents there is no correlation among the two variables. Once the covariance matrix is formulated the next task is to compute the eigen values and vectors.

Eigen vectors are always non-zero vectors and each individual vector represents the principal component. Eigen values are of scalar type and eigen vector with the highest eigen value represents the first principal component, and further principal components are decided as per eigen values in the descending order.

$$\begin{pmatrix} V(y_1, y_2), C(y_1, y_2) \dots \dots C(y_1, y_z) \\ C(y_2, y_1), V(y_2, y_2) \dots \dots C(y_2, y_z) \\ \text{-----} \\ C(y_z, y_2), C(y_z, y_2) \dots \dots V(y_z, y_z) \end{pmatrix}. \quad (3)$$

The major task while calculating eigen vectors and values is to maximize the variance. It can be done through stacking  $p$  data vectors into  $p \times q$  matrix,  $a$ , then, the projections are given by  $ab$ , which is an  $p \times 1$  matrix [22]. Variance could be computed as:

$$\begin{aligned} \sigma_{\vec{b}}^2 \frac{1}{p} * \left( \sum_i (\vec{a}_i \cdot \vec{b})^2 \right) &= \frac{1}{p} * ((\mathbf{ab})^T (\mathbf{ab})) \\ &= \frac{1}{p} * (\mathbf{b}^T \mathbf{a}^T \mathbf{a} \mathbf{b}) \\ &= \mathbf{b}^T * \frac{\mathbf{a}^T \mathbf{a}}{p} * \mathbf{b} = \mathbf{b}^T \mathbf{v} \mathbf{b}. \end{aligned} \quad (4)$$

It is desirable to select a unit vector  $\vec{b}$  so that there is maximum  $\sigma_{\vec{b}}^2$ . For doing so, unit vectors must be analyzed so that maximization is constrained. After successfully implementing the above strategy, the maximum variance

can be obtained, and it can further help in finding the best possible eigen vectors for computing principal components. We can conclude that there is  $v$  matrix with dimensions  $k \times k$  where  $k$  denotes different eigen vectors. Eigen vectors are always orthogonal to each other. These computed eigen vectors are also known as principal components of data. The first principal component is one, which is having the maximum data variance, where the next set of principal components are decided as per next levels of data variance. Algorithm 1 is specifying the overall process for dimensionality reduction and extracting essential features from the entire set.

Here,  $\Sigma$   $\mathbf{a}$  denotes the covariance matrix from which principal components will be generated, and overall data  $Z$  will be projected into PCA subspace after calculating eigen values and vectors.

**3.7. Task Classification Using Support Vector Machines (SVM, RBF-Kernel Based).** Task classification is one of the major steps involved which decides on whether the submitted request in the form of a series of data values is critical or non-critical. For doing so, supervised machine learning algorithm is implemented with the help of radial basis function kernel. SVM can be used for both classification and regression analysis. Here, it is implemented for the classification problem. SVM works with the principal of finding hyperplane that can isolate the features into classes. Kernels can play a major role in high dimensional data. Here, in the proposed work, data of numerous features is collected which is used for further system training. It becomes necessary to implement the kernel trick for getting a better system accuracy. Radial basis function kernel is also known as Gaussian-based kernel which is primarily used for non-linear data. Following equation determines the mathematical model of RBF-based kernel.

$$A(P, P') = \exp(-\text{gamma} * (\|P - P'\|)^2) \quad (5)$$

Here, gamma value can vary from 0 to 1. Correct usage of gamma value plays an important role. Values such as 0.1 is the most preferred value to avoid underfitting and overfitting of the data. A dataset with  $p$  no. of samples is denoted as:  $\{\vec{Z}_j, \mathbf{n}_j\}, j=1, 2, \dots, p$  with  $\vec{Z}_j \in \{\mathbf{0}, \mathbf{1}\}^x$  exhibits sample data and  $\mathbf{n}_j \in \{+1, -1\}$  denotes class label of given sample. For a testing sample  $Z$ , RBF-kernel-based SVM classification formula is as follows:

$$F(Z) = \text{signfunc}\left(c + \sum_{j=1}^p \gamma_j \mathbf{n}_j \mathbf{K}(\vec{Z}_j, Z)\right). \quad (6)$$

In the given equation  $\gamma_j$  and  $c$  are knowledge factors of SVM and given  $\text{signfunc}()$  is the sign function whereas  $\mathbf{K}(\vec{Z}_j, Z)$  represents a kernel function [24]. Projection of RBF kernel is also done into infinite dimensions, such as:

$$A(P, P') = \langle \boldsymbol{\varphi}(P), \boldsymbol{\varphi}(P') \rangle. \quad (7)$$

Here,  $\boldsymbol{\varphi}$ , is a type of function, that projects vectors  $y$  into another vector space. This function also calculates inner-product between two vectors. It can also project vectors into infinite dimension and proof for the same is also provided as per following:

$$\boldsymbol{\varphi}_{\text{IRBF}} = \mathbf{K} \mathbf{n} \longrightarrow \mathbf{K}^{\infty}. \quad (8)$$

Consider gamma value in equation (5) to be  $\frac{1}{2}$ ,

$$\begin{aligned} AIRBF(P, P') &= \exp\left[-\frac{1}{2}\|P - P'\|^2\right] \\ &= \exp\left[-\frac{1}{2}\langle P - P', P - P' \rangle\right] \\ &= \exp\left[-1/2(\langle P, P - P' \rangle - \langle P, P - P' \rangle)\right] \\ &= \exp\left[-\frac{1}{2}(\|P\|^2 + \|P'\|^2 - 2\langle P, P' \rangle)\right] \\ &= \exp\left[-\frac{1}{2}(\|P\|^2 - \frac{1}{2}\|P'\|^2)\right] \exp\left[-\frac{1}{2} - 2\langle P, P' \rangle\right] \\ &= G e^{\langle P, P' \rangle}, \text{ Here } G \text{ is constant} \\ &= G \sum_{n=0}^{\infty} \langle P, P' \rangle^n / n! \\ &= G \sum_{n=0}^{\infty} A_{\text{poly}(n)}(P, P') / n!. \end{aligned} \quad (9)$$

Equation (9) proves that RBF kernel is formed by performing infinite sum over polynomial kernels. Consider the sum of two kernels, such as:

$$\begin{aligned} Az(P, P') &:= Ax(P, P') + Ay(P, P'), \\ \boldsymbol{\varphi}_z(P) &= (\boldsymbol{\varphi}_x(P) + \boldsymbol{\varphi}_y(P)). \end{aligned} \quad (10)$$

Here,  $\boldsymbol{\varphi}_z(P)$  is a tuple, where the first element is the vector  $\boldsymbol{\varphi}_x(P)$ , and the other element is  $\boldsymbol{\varphi}_y(P)$ . Inner product is given by:

$$\langle \boldsymbol{\varphi}_z(P) \boldsymbol{\varphi}_z(P') \rangle := \langle \boldsymbol{\varphi}_x(P) \boldsymbol{\varphi}_x(P') \rangle + \langle \boldsymbol{\varphi}_y(P), \boldsymbol{\varphi}_y(P') \rangle. \quad (11)$$

In general, it is given as:

$$\begin{aligned} \langle \boldsymbol{\varphi}_z(P), \boldsymbol{\varphi}_z(P') \rangle &:= \sum_i^{\text{di mension}(x)} \boldsymbol{\varphi}_{x,i}(P) \boldsymbol{\varphi}_{x,i}(P') \\ &\quad + \sum_j^{\text{di mension}(x)} \boldsymbol{\varphi}_{y,j}(P) \boldsymbol{\varphi}_{y,j}(P'). \\ &= \sum_i^{\text{di mension}(x) + \text{di mension}(y)} \boldsymbol{\varphi}_{z,i}(P) \boldsymbol{\varphi}_{z,i}(P') \end{aligned} \quad (12)$$

Finally, this is the projection into a vector space with infinite dimension.



module PCA(Z)

- (i) Compute sample mean and sample covariance matrix:  $(\mu_a = (1/p) * \sum_{i=1}^p \mathbf{a}_i), \sum_a = (1/(p-1)) * (\sum_{i=1}^p (\mathbf{a}_i - \mu_a)(\mathbf{a}_i - \mu_a))$
  - (ii) Compute eigen values and eigen vectors of  $\sum_a$
  - (iii) Formulate transformation matrix  $S = [t_1, t_2, \dots, t_y]$  with  $y$  eigen vectors in association with  $y$  largest eigen values.
  - (iv) Now, do the projection of overall data  $Z$  into the subspace of PCA such as:  $r_i = Sa_i$ , for  $i = 1, 2, \dots, p$ .
  - (v) Return  $r$
- End module

ALGORITHM 1: Principal component analysis for dimensionality reduction [23].

**3.8. Workload Optimization Using Genetic Algorithm.** Due to heavy network traffic, there is a certain chance that a couple of fog nodes may get overloaded. So, it becomes necessary to perform the optimization after certain interval of time. Workload optimization procedure will try to locate the idle fog nodes. Once the idle fog nodes are identified, fog offloading could be implemented and various requests could be migrated from overloaded foglet. For doing so, genetic algorithm (GA) is implemented as a part of the working methodology. Genetic algorithm works on the principal of natural evolution. It works by choosing the fittest population from the given set. Firstly, the parents with the best surviving capabilities are identified, and later, their offspring will come, which could be better as compared to their parents. This process keeps on happening, and at the end, the best population set is achieved. It works in five phases such as: (a) selection of initial population, (b) fitness function, (c) selection, (d) crossover, and (e) mutation [25].

- (a) So, it starts by choosing the initial populations where genes which characterizes an individual are selected to form chromosomes, and by doing so, the final population is decided. In the proposed system also, initial set of fog nodes are decided, such as: F1, F2, F3, F3...
- (b) The next step is to choose a fitness function which determines the ability of an individual to survive with the other set of individuals available in the system. So, for doing so, fitness score will be calculated for the individual. Same scenario is simulated for the available fog nodes, and their fitness scores are calculated
- (c) After selecting the fittest individuals, routines will be designed to pass on the genes (or acquired features) into the next generations to obtain the optimized results
- (d) next step is to perform crossover, where crossover points are decided up to which parents can keep on sharing genes and offspring is generated.
- (e) last step is to perform mutation where certain changes are performed in the genes of offspring to bring versatility in the generated population.

All steps mentioned above use the concept of survival of the fittest. Genetic operators will be implemented on the list of individuals to generate new population. Following is the

mathematical representation of various steps involved in GA optimization process:

Consider initial population(Z) of  $x$  chromosomes. The first task is to calculate fitness value of every chromosome. Consider there are two chromosomes CH1 and CH2 chosen from the overall population as per their respective fitness values. There is a crossover probability associated with every chromosome, and same is applied on CH1 and CH2 to produce new offspring, such as Osp. Once offspring (Osp) is generated, it is correlated with mutation probability to generate new offspring (NOsp). This newly generated offspring will be a part of the new population. The process of selection, crossover, and mutation will be executed again and again until unless the new population is completely formulated. The existing search process of GA is continued to achieve the optimal solution. The global search capabilities of GA are better as it can assess numerous individuals, and hence, a variety of optimal solutions could be generated. Crossover is the vital process which will generate the offspring, and it is denoted as:

$$\mathbf{X} = \frac{(\mathbf{P} + 2\sqrt{s})}{3\mathbf{P}} \quad (13)$$

Here,  $s$  is the number of generations,  $P$  is the total number of evolutionary generation sets by population.  $X$  is dynamic in nature, and it can vary, as the number of evolutionary generations can also change with constant intervals of time. Complete specification of the optimization process is denoted by Algorithm 2.

Following steps are giving the complete details of fitness function and supporting equations with respect to modified GA for workload optimization:

- (a) Total no. of individuals at every time step  $x$  and can satisfy  $Z$  are given as:

$$\mathbf{I}_{z,x} = |Y_x \cap Z|. \quad (14)$$

- (b) Following expression provides observed average fitness at time  $x$ :

$$\bar{\mathbf{a}}(x) = \frac{1}{n} * \sum_{i=1}^n \mathbf{a}(\mathbf{q}_i, x). \quad (15)$$

- (c) Term  $\bar{\mathbf{a}}(Z,x)$  means schema's observed average fitness, where schema is denoted by  $Z$ , at time  $x$ , such as:

```

Input:
  Initial population size: x
  Consider total no. of iterations: y
Output:
  Optimized solution(Os)
Start
  Formulate initial population of x chromosomes, Oi (i = 1, 2, . . . , n)
  Initialize looping variable, i = 0
  Calculate fitness value of all chromosomes
  while (i < y)
    select pair (CH1, CH2) as per fitness value
    Implement crossover on (CH1, CH2) pair using crossover probability
    Implement mutation on offspring (Osp) using mutation probability
    Overwriting existing population with new population
    Increment looping variable i by 1
  end while
  return optimized solution (Os)
Stop

```

ALGORITHM 2: Genetic algorithm for workload optimization

$$\overline{\mathbf{a}}(Z, \mathbf{x}) = \frac{1}{|Z, \mathbf{X}|} * \sum_{i \in \{j\} | q_{j, x \in z}} \mathbf{a}(\mathbf{q}_i, \mathbf{x}). \quad (16)$$

Following inequality holds for every schema Z

$$E[|Z, \mathbf{x} + 1|] > = |Z, \mathbf{x}| \cdot \overline{\mathbf{a}}(Z, \mathbf{x}) / \overline{\mathbf{a}}(\mathbf{x}) \cdot \frac{(1 - \text{Pr}_c - \gamma(Z))}{\mathbf{n} - 1} \cdot (1 - \text{Pr}_m)^{\alpha(H)}. \quad (17)$$

Probability of choosing an individual, satisfying Z is:

$$\left( \frac{\sum_{i \in \{j\} | q_{j, x \in z}} \mathbf{a}(\mathbf{q}_i, \mathbf{x})}{\sum_{i=1}^{\mathbf{n}} \mathbf{a}(\mathbf{q}_i, \mathbf{x})} \right). \quad (18)$$

Given that the probability is a constant and remains the same in complete iterations of the loop. To obtain desired number of selected individuals satisfying Z (with sample amount: n),

$$\begin{aligned} \mathbf{n} \cdot \sum_{i \in \{j\} | q_{j, x \in z}} \frac{\mathbf{a}(\mathbf{q}_i, \mathbf{x})}{\sum_{i=1}^{\mathbf{n}} \mathbf{a}(\mathbf{q}_i, \mathbf{x})} &= \mathbf{n} \cdot |Z, \mathbf{x}| / |Z, \mathbf{x}| \cdot \sum_{i \in \{j\} | q_{j, x \in z}} \frac{\mathbf{a}(\mathbf{q}_i, \mathbf{x})}{\sum_{i=1}^{\mathbf{n}} \mathbf{a}(\mathbf{q}_i, \mathbf{x})} \\ &= |Z, \mathbf{x}| \cdot \overline{\mathbf{a}}(Z, \mathbf{x}) / \overline{\mathbf{a}}(\mathbf{x}) \end{aligned} \quad (19)$$

If there are two individuals crossed, then, the probability that the cross site is selected within the given length of Z is:

$$\alpha(Z / \mathbf{n} - 1) \quad (20)$$

Survival probability Pr<sub>s</sub> of Z:

$$\alpha(Z / \mathbf{n} - 1) \quad (21)$$

Here, Pr<sub>c</sub> is the crossover probability.

After executing selection and crossover, we can compute no. of strings satisfying Z after crossover using

$$\begin{aligned} \overline{\mathbf{a}}(Z, \mathbf{x}) / \overline{\mathbf{a}}(\mathbf{x}) \cdot |Z, \mathbf{x}| \cdot \text{Pr}_s > \\ = \overline{\mathbf{a}}(Z, \mathbf{x}) / \overline{\mathbf{a}}(\mathbf{x}) \cdot |Z, \mathbf{x}| \cdot (1 - \text{Pr}_c \cdot \alpha(Z / \mathbf{n} - 1)). \end{aligned} \quad (22)$$

Probability that all specifications of Z are unaffected by mutation is:

$$(1 - \text{Pr}_m)^{\delta(Z)}. \quad (23)$$

## 4. Results and Discussion

The proposed methodology was implemented in the simulation environment and various experimental settings were performed. Table 3 demonstrates the hardware and software needs for conducting the experiment.

*4.1. Software and Hardware Description.* The proposed fog-computing-based smart healthcare system is simulated with the help of iFogSim simulation tool. This tool is perfect for modelling the proposed work with the support of all possible topologies and scenarios.

Patient data is managed on the MySQL, and the same data is used to train the prescribed classifier. The data of 100 users was considered in the simulation where data is collected through various biomedical sensors.

*4.2. Experimental Settings.* Table 4 shows the experimental settings for implementing the proposed smart healthcare system. Figure 3 shows the simulation snapshot of 4 fog nodes. Simulation is extended up to 10 fog nodes.

*4.3. Comparative Analysis on the Basis of Various Parameters.* Various parameters were chosen for the comparative analysis such as latency, execution time, accuracy, and system stability. Various subsections following this section performed the required analysis on the acquired results. For

TABLE 3: Software and hardware description.

Software/Hardware	Description
Simulation tool	iFogSim
Simulator Version	3.0.3
Operating system	Windows 10
Programming Language	Java
JDK version	Java SE 12
IDE	Eclipse IDE 2021-03
Database	MySQL 8.0.24

TABLE 4: Experimental settings.

Parameter	Sub-parameter	Value
Number of users		100
Count of biomedical sensors		10
Count of fog nodes		10
Fog node configuration	Storage	2 GB
	Bandwidth	1500 KBs
	Resource cost	3.0-2.5
	Memory cost	0.5-0.4

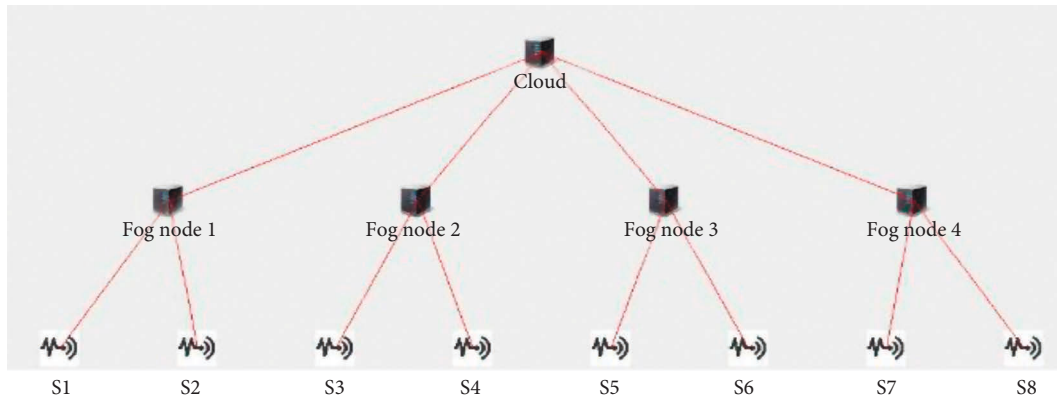


FIGURE 3: iFogSim simulation snapshot.

comparative analysis, various existing systems are chosen from the literature survey. Fog-BBN [15], Fog-Smart Office [16], Fog-CPE [20], Tri-Fog Health [21].

**4.3.1. Latency Analysis.** Latency is defined as the time taken for responding to the submitted request on the fog node. It is denoted by the following equation which includes the propagation and execution time.

$$l = t_p + e. \quad (24)$$

Here,  $t_p$  represents the propagation time, and  $e$  is the execution time.

Figure 4 shows the latency comparison of various existing fog-based systems. Due to the preliminary steps performed before going into the fog layer, lesser latency could be observed in the proposed work. Workload optimization has significantly worked in the favour of reducing latency, consequently improving the overall system performance. State of the art methods are lacking load optimization, hence more delay can be observed.

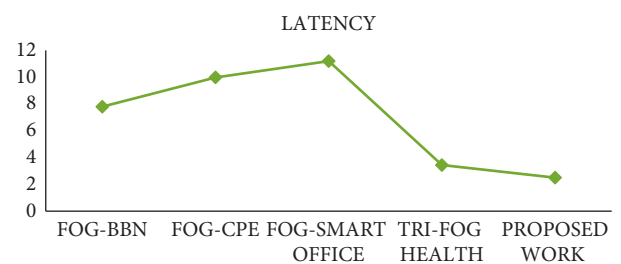


FIGURE 4: Latency comparison.

**4.3.2. Analysis of Execution Time.** Execution time (in milliseconds) with respect to the healthcare system is the total time taken to decide whether the submitted request is critical or non-critical. Based on the criticality of the submitted request, it will be forwarded for next action. Figure 5 shows the comparison of execution time for various existing fog-based systems. Optimization techniques, classification of task criticality, and noise removal process implemented in the fog layer have contributed to speed up the execution process. IoT applications require faster task execution, so

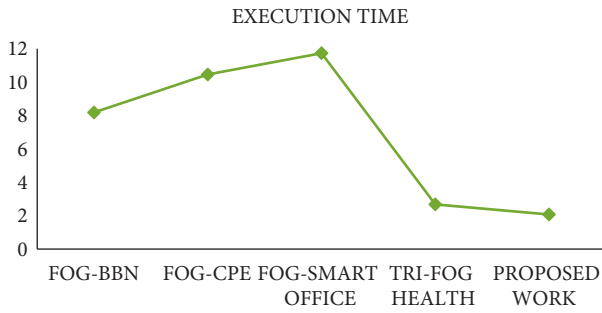


FIGURE 5: Comparison of execution time.

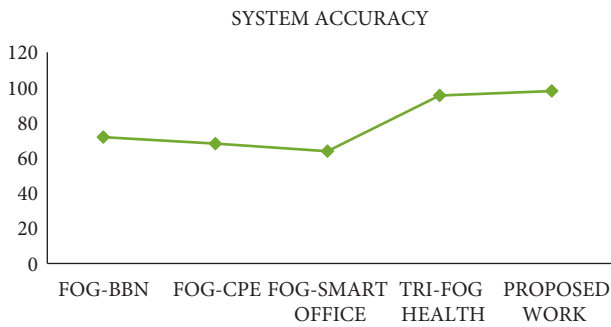


FIGURE 6: Overall system accuracy.

that speedy results could be extracted for further processing. So, the proposed work has catered this requirement.

**4.3.3. Overall System Accuracy.** Comparison on the basis of overall system accuracy (in terms of %) was also performed. Accuracy is measured in terms of valid requests submitted to the fog nodes, true alerts for the critical tasks, and cleaned data for training the system. Keeping all these factors in view, Figure 6 shows the accuracy comparison for various existing fog-based systems, and the proposed work is overpowering all of them. State of the art methods are not performing sensor data cleaning through some preprocessing technique, hence suffering a lower system accuracy. Data analysis is another important factor here which can be managed properly. The proposed methodology is taking care of these aspects efficiently.

**4.3.4. System Stability.** This parameter determines the scalability of the system and its ability to handle increasing number of users. Simulations were conducted on a large number of users (i.e., 100), and stability factor was determined for the same. Figure 7 highlights the capability of the proposed work in terms of handling large number of users. Improvement in the other three performance metrics has certainly made sure that the system can work appropriately with the increase in the number of users.

**4.4. Summary of Results.** Comparative analysis mentioned in section 4.3 is clearly depicting the efficiency of proposed work in the field of healthcare systems. Table 5 depicts the

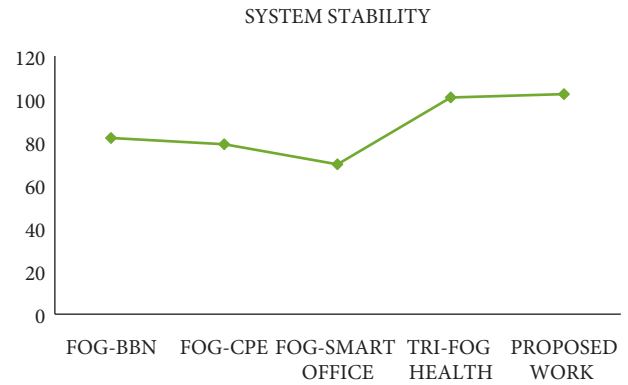


FIGURE 7: Comparison of system stability.

TABLE 5: Results summary.

Parameter	Fog-BBN	Fog-CPE	Fog-Smart Office	Tri-Fog Health	Proposed Work
Latency (ms)	7.8	9.98	11.2	3.44	2.5
Execution Time (ms)	7.8	9.98	11.2	2.52	1.94
System Accuracy (%)	72.4	68.8	64.6	95.44	97.9
System Stability (%)	78.8	76	67	97	98.5

summary of the results gathered after performing the simulations.

Proposed work has shown considerable amount of improvement as compared to the recent and optimized literature work, i.e., Tri-Fog Health [21]. There is a percentage decrease of 27% in latency metric. Execution time drops by 23%, which shows the utility of proposed work. System accuracy increases by a factor of 2.57% and system stability is improved with a measure of 1.54%

## 5. Conclusions and Future Scope

In the proposed work, a novel framework for smart healthcare system is presented. Fog computing paradigm works closer to the IoT layer, so that processing could be done at the edges only and speedy response could be given. The presented methodology is implemented with the same concept. Fog layer could experience some delays and sometimes lack in the performance, as resource constraints are there. Presented framework tries to work on the same aspects and noticeable results were obtained. Variety of activities were also performed to achieve refinements in the results. This work could be extended to work in various other domains such as: Smart cities, Smart Homes, Smart Grids. These domains involve such activities which are time critical, and significant delays are not tolerable. Along with different domains, work could be done to introduce deep learning concepts for imparting better learning to the system. Scalability is another factor which is required to be improved, as in real-time scenarios, we may get more

number of users, and a large amount of data is required to be processed.

### Data Availability

The data presented in this study are available on request from the corresponding author.

### Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

### Acknowledgments

The authors would like to thank the research Chair of Prince Faisal for Artificial Intelligence (CPFAI) for supporting this research work.

### References

- [1] M. Ghobaei-Arani, A. Souiri, and A. A. Rahmanian, "Resource management approaches in fog computing: a comprehensive review," *Journal of Grid Computing 2019*, vol. 18, pp. 1–42, 2019.
- [2] C.-H. Hong and B. Varghese, "Resource management in fog/edge computing," *ACM Computing Surveys*, vol. 52, pp. 1–37, 2019.
- [3] H. A. Khattak, H. Arshad, S. Ul Islam et al., "Utilization and load balancing in fog servers for health applications," *EURASIP Journal on Wireless Communications and Networking*, vol. 91, 2019.
- [4] M. H. Kashani, A. Ahmad, and E. Mahdipour, "Load balancing mechanisms in fog computing: a systematic review," 2020, <https://arxiv.org/abs/2011.14706>.
- [5] S. Tuli, N. Basumatary, S. S. Gill et al., "Healthfog: an ensemble deep learning based smart healthcare system for automatic diagnosis of heart diseases in integrated IoT and fog computing environments," *Future Generation Computer Systems*, vol. 104, pp. 187–200, 2020.
- [6] K. Gai, Z. Lu, M. Qiu, and L. Zhu, "Toward smart treatment management for personalized healthcare," *IEEE Network*, vol. 33, pp. 30–36, 2019.
- [7] M. K. Hassan, A. I. E. Desouky, M. M. Badawy, A. M. Sarhan, M. Elhoseny, and M. Gunasekaran, "EoT-driven hybrid ambient assisted living framework with naïve bayes–firefly algorithm," *Neural Computing and Applications*, vol. 31, pp. 1275–1300, 2018.
- [8] F. Sun, W. Zang, R. Gravina, G. Fortino, and Y. Li, "Gait-based identification for elderly users in wearable healthcare systems," *Information Fusion*, vol. 53, pp. 134–144, 2020.
- [9] A. S. Albahri, O. S. Albahri, A. A. Zaidan et al., A. H. Alamoodi, O. Enaizan, S. Nidhal et al., "Based multiple heterogeneous wearable sensors: a smart real-time health monitoring structured for hospitals distributor," *IEEE Access*, vol. 7, pp. 37269–37323, 2019.
- [10] A. Anagaw and Y.-L. Chang, "A new complement naïve Bayesian approach for biomedical data classification," *Journal of Ambient Intelligence and Humanized Computing*, vol. 10, pp. 3889–3897, 2018.
- [11] Y. Liu, L. Zhang, Y. Yang et al., "A novel cloud-based framework for the elderly healthcare services using digital twin," *IEEE Access*, vol. 7, pp. 49088–49101, 2019.
- [12] H. Zhang, J. Liu, and N. Kato, "Threshold tuning-based wearable sensor fault detection for reliable medical monitoring using bayesian network model," *IEEE Systems Journal*, vol. 12, pp. 1886–1896, 2016.
- [13] I. Azimi, J. Takalo-Mattila, A. Anzanpour, A. M. Rahmani, J.-P. Soininen, and P. Liljeberg, "Empowering healthcare IoT systems with hierarchical edge-based deep learning," in *Proceedings of the 2018 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies*, New York, NY, USA, September 2018.
- [14] B. Guo, Y. Ma, J. Yang, and Z. Wang, "Smart healthcare system based on cloud-internet of things and deep learning," *Journal of Healthcare Engineering*, vol. 2021, Article ID 4109102, 2021.
- [15] P. Verma and S. K. Sood, "Fog assisted-IoT enabled patient health monitoring in smart homes," *IEEE Internet of Things Journal*, vol. 5, pp. 1789–1796, 2018.
- [16] M. Bhatia and S. K. Sood, "Exploring temporal analytics in fog-cloud architecture for smart office healthcare," *Mobile Networks and Applications*, vol. 24, pp. 1392–1410, 2018.
- [17] S. R. Hassan, I. Ahmad, S. Ahmad, A. Alfaify, and M. Shafiq, "Remote pain monitoring using fog computing for e-healthcare: an efficient architecture," *Sensors*, vol. 206574 pages, 2020.
- [18] S. R. Hassan, I. Ahmad, J. Nebhen, A. U. Rehman, M. Shafiq, and J. G. Choi, "Design of latency-aware IoT modules in heterogeneous fog-cloud computing networks," *CMC-COMPUTERS MATERIALS & CONTINUA*, vol. 70, no. 3, pp. 6057–6072, 2022.
- [19] S. R. Hassan, I. Ahmad, A. U. Rehman, S. Hussen, and H. Hamam, "Design of resource-aware load allocation for heterogeneous fog computing environments," *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 3543640, 2022.
- [20] R. Dautov, S. Distefano, and R. Buyya, "Hierarchical data fusion for Smart Healthcare," *Journal of Big Data*, vol. 6, 2019.
- [21] M. Ijaz, G. Li, H. Wang et al., "Intelligent fog-enabled smart healthcare system for wearable physiological parameter detection," *Electronics*, vol. 9, 2020.
- [22] H. Abdi and L. J. Williams, "Principal component analysis," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 2, pp. 433–459, 2010.
- [23] A. Tharwat, "Principal component analysis-a tutorial," *International Journal of Applied Pattern Recognition*, vol. 3, no. 3, pp. 197–240, 2016.
- [24] Q. Liu, C. Chen, Y. Zhang, and Z. Hu, "Feature selection for support vector machines with RBF kernel," *Artificial Intelligence Review*, vol. 36, pp. 99–115, 2011.
- [25] M. Kumar, M. Husain, N. Upreti, and D. Gupta, "Genetic algorithm: review and application," *SSRN Electronic Journal*, 2010.

## Research Article

# Application of Deep Learning in College Physical Education Design under Flipped Classroom

Jun Huang  and Dian Yu 

*School of Competitive Sports, Shandong Sport University, Rizhao 276827, Shandong, China*

Correspondence should be addressed to Dian Yu; [yudian@sdpei.edu.cn](mailto:yudian@sdpei.edu.cn)

Received 14 June 2022; Revised 27 July 2022; Accepted 10 August 2022; Published 16 September 2022

Academic Editor: Inam Ullah

Copyright © 2022 Jun Huang and Dian Yu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the development of information technology, teaching reform has also undergone major changes. The traditional college physical education teaching method cannot meet the needs of the majority of students, and the physical education teaching mode continues to be reformed. Microcourse is the most intuitive form of deep integration of information technology and physical education. From the perspective of the flipped classroom (FC), the physical education model has gradually changed from teacher centered to student centered. Deep learning (DL) emphasizes that learners have the ability to actively construct knowledge, effectively transfer knowledge, and solve real problems. This design applies DL and convolutional neural network to the teaching design of physical gymnastics in colleges and universities. The application of the DL teaching model based on FC in the microcourse teaching of gymnastics in colleges and universities is studied and evaluated. The results show that the current utilization of microcourse teaching resources is too low. Interest-oriented teaching microcourses cannot improve students' interests. The proportion of students who are interested is relatively small, and more than 50% of students are not interested. Teachers generally believe that the current gymnastics microcourse needs further optimization and improvement. The poor quality of microvideos and the lack of supervision and reward mechanism in the course are the main reasons for the insufficient students' interest. The complexity of the videos and the liveliness of the discussions are the main problems of low resource utilization. The student's interest in learning is greatly improved after the application of the designed model, and the proportion increases to 82.4%. The effect on ordinary college students is the most obvious, and the effect of microvideo learning has been significantly promoted. Design mode has the most obvious improvement in improving learning efficiency and autonomous learning ability. The improvement of learning ability has increased from 18% to 72%, and the improvement of learning efficiency has increased from 39% to 82%. Meanwhile, students' interest in learning is stimulated, and the utilization of resources is improved.

## 1. Introduction

With social progress and technological development, China's educational reform level has also been greatly improved. Physical education has also been valued by colleges and all sectors of society, but the final teaching effect is not very obvious. Deepening teaching reform is crucial. The reform of teaching methods is an important manifestation of teaching reform. New ideas have been brought to the innovation of physical education teaching in colleges and universities with the continuous deepening of the flipped classroom (FC) education model and the gradual use of

information-based teaching methods. In addition, it brings the possibility to improve the quality of physical education teaching in colleges and universities. The FC has played a huge role in promoting the reform of physical education teaching in colleges and universities [1].

In today's society, scholars have also conducted various researches on physical education in colleges and universities. Some researchers have deeply discussed how to apply FC to taekwondo teaching in colleges and universities. They have studied whether FC can help improve the taekwondo skills of college students through educational practice and questionnaire methods [2]. Hinojo Lucena et al. used the method

of literature analysis and case analysis to take “FC” as the research object and focused on analyzing its application strategies in physical education [3]. Julia and Marco analyzed the basic theoretical knowledge of public sports as the teaching content of the public sports massive open online courses (MOOCs) using the literature method and questionnaire survey method. They compared the implementation effect of MOOC technology [4]. Fang et al. combined the FC teaching mode with physical education teaching based on knowledge construction and made an objective teaching evaluation of this teaching mode [5]. Liao proposed the practical application of the FC model in the teaching of sports table tennis in colleges and universities under the background of MOOCs [6]. Students are mostly in a passive acceptance state in the traditional lecture-style teaching process. Simple memory and repetitive training hinder the transfer of learner knowledge [7]. Deep learning (DL) advocates for learners to actively apply what they have learned to solve complex problems in reality. It helps to enhance critical thinking and knowledge construction [8]. Tufail et al. applied a DL 3D convolutional neural network (CNN) architecture to disease diagnosis [9]. Tufail et al. conducted a taxonomic study of the endangered jujube species in the European habitat ecosystem based on a DL architecture model [10]. Khan et al. applied DL models to the multiclass classification problem of modulated signals in intelligent communication systems [11]. Tufail et al. applied DL to help determine optimal model designs for cancer diagnosis and prediction tasks [12]. The integration of DL and teaching has become a research hotspot in education. Shuo and Ming [13] combined DL with smart classrooms and designed a “3 + 1” smart teaching model. Jan et al. [14] summarized the process of educational big data mining based on the characteristics of educational big data mining under DL. Through literature research, it is found that DL models are widely studied in medical care, ecosystems, and communications and are also infiltrated in teaching. Research on physical education teaching in colleges and universities is scarce.

Based on this, under the concept of FC, this study innovatively combines the CNN model in the DL algorithm with the teaching design of college sports gymnastics microcourse. The application of the teaching mode is investigated through a questionnaire survey, and the application effect is further analyzed. This design aims to improve the teaching effect of gymnastics microcourse, stimulate learning interest, and promote the wide application of microcourse resources by changing the gymnastics teaching mode. Although there are some limitations, it is expected that the improvement of the design concept and teaching mode can provide a reference for the subsequent reform and construction of college physical education.

## 2. Methods

**2.1. FC.** FC is the product of the combination of network and face-to-face teaching platforms and is a new teaching model formed under the support of modern information technology [15]. This learning model can be widely applied to various educational and teaching processes. FC frees

teachers from tedious teaching procedures, leaving time to focus on students. Teachers establish a close relationship with students to enhance students’ autonomous learning motivation [16]. FC innovates the traditional education model, and the teaching model has many characteristics, such as innovative teaching mode, student-centered differentiated teaching mode, personalized blended learning mode, and dynamic classroom [17]. In this education model, students have changed from passive to active, and their learning methods have also changed to a model of preclass learning and classroom inquiry. The teaching method is not limited by the location, which is convenient for students to learn [18]. Taking student learning as the center can continuously stimulate students’ subject consciousness and cognitive ability. Besides, it stimulates the creativity and participation of students.

**2.2. DL Theory.** DL is a machine learning method with multiple neural network layers, which has a wide range of applications in many fields. DL has been applied in graph analysis, data processing, and public data prediction and has powerful feature information extraction and learning capabilities [19, 20]. Learners can critically learn new knowledge and update existing cognitive structures while transferring knowledge to new situations [21]. DL extracts features layer by layer by mining the underlying feature distribution of the data and transforms low-level feature data into high-dimensional features [22]. DL consists of three layers, namely, input layer, hidden layer, and output layer.

DL can use standardization, normalization, and other operations for data preprocessing to make the data meet the training needs [23]. Normalized data are handled by scaling with the two maxima:

$$X_i = \frac{X_i - \min(X)}{\max(X) - \min(X)}. \quad (1)$$

Normalization is a transformation for each feature. In equation (1),  $\max(X)$  represents the maximum value of a feature and  $\min(X)$  represents the minimum value of a feature. The value range of all data transformations is [0, 1].

Normalization is to make the sample size have a uniform standard when calculating the similarity by dot product or kernel function. Each sample is transformed, and the original data are transformed into a unit vector. The conversion process is shown as follows:

$$x = \frac{x}{\sqrt{\sum_j^n X[j]^2}}. \quad (2)$$

In equation (2),  $n$  is the number of sample features and  $X[j]$  is the  $j$ th feature of sample  $x$ . The function of normalization is to make the sample vectors have a uniform standard when similarity is calculated by dot multiplication or kernel function.

CNN is one of the typical algorithms in DL. Partial connections are used between the internal layers of the CNN. Weight sharing can be achieved between neurons in

the same layer. CNN model parameters are reduced, which has obvious advantages in speech recognition and image recognition tasks. A CNN consists of an input layer, a convolutional layer, a pooling layer, a fully connected layer, and an output layer [24]. The DL model is shown in Figure 1.

From Figure 1, the input layer is responsible for data input and further preprocesses the input data to pave the way for the next step of CNN convergence. Convolutional layers use convolution to extract multiple features of input data or images. When the convolution kernel performs convolution processing, it scans the input data according to step size, performs matrix multiplication and summation in the convolution kernel, and adds the deviation:

$$\begin{aligned} Z^{l+1}(i, j) &= [Z^l \otimes w^{l+1}](i, j) + b \\ &= \sum_{k=1}^{K_l} \sum_{x=1}^f \sum_{y=1}^f [Z_k^l(s_0 i + x, s_0 i + y) w_k^{l+1}(x, y)] + b. \end{aligned} \quad (3)$$

In equation (3),  $b$  is the deviation.  $Z^l$  and  $Z^{l+1}$  are the convolution input and output of the  $l + 1$ th layer.  $Z(i, j)$  is the pixel of the feature map, and  $K$  is the number of channels of the feature map.  $f$  and  $s_0$  are the size of the convolution kernel and the convolution step size, respectively.

The value range of pixel  $(i, j)$  is shown in the following equation:

$$(i, j) \in \{0, 1, \dots, L_{l+1}\}, \quad (4)$$

$$L_{l+1} = \frac{L_1 + 2p - f}{s_0} + 1. \quad (5)$$

In equations (4) and (5),  $p$  is padding and  $L_{l+1}$  is the size of  $Z^{l+1}$ .

The pooling layer compresses the convolutional features and uses the pooling function in the pooling layer to count the value of each area in the feature map, including pooling size, step size, and padding. The pooling method is shown in the following equation:

$$A_k^l(i, j) = \left[ \sum_{x=1}^f \sum_{y=1}^f A_k^l(s_0 i + x, s_0 i + y)^p \right]^{1/p}. \quad (6)$$

In equation (6), the pixel  $(i, j)$  is the same as the convolutional layer and  $p$  is the pooling parameter.

After the image passes through the pooling layer, it is flattened and passed to the fully connected layer. At this time, the image loses its spatial characteristics and is expanded into the form of a vector.

### 3. Application of the DL Teaching Mode Based on FC in Microcourse Teaching of Physical Gymnastics in Colleges and Universities

**3.1. Microcourse.** “Physical education professional gymnastics microclass” focuses on the teaching goals and difficulties of gymnastics courses. It takes teaching videos as the main carrier, emphasizes active learning, and carries out

online network teaching activities in network media. In general, microcourses are presented in the form of microvideos [25]. As a kind of course resource, the specific content of the microcourse is shown in Figure 2.

In Figure 2, the teaching resources of microcourses mainly include microtargets, microvideos, microteaching plans, microexercises, microdiscussions, and micro-evaluations. Various teaching resources of microcourse serve the teaching process and students’ autonomous learning process. The teaching application mode of microcourse is divided into three types, namely, flipped teaching, classroom differentiated teaching, and after-school tutoring and answering application mode. Among them, the flipped teaching application mode is the most commonly used teaching mode in microcourses. The microcourses under its teaching application mode can be placed before the class or during the class, forming a model of “learning first and teaching later.”

Gymnastics is a compulsory course for physical education and training, covering the characteristics of physical education and being representative. Physical education professional gymnastics microcourse is taken as the research object. The microcourse teaching mode of gymnastics in colleges and universities is designed using the in-depth learning route. The design content includes seven parts, which are determining teaching objectives and content, preassessment and analysis, interest-oriented stimulation of students’ interest in learning, learning with microcourse resources, learning task transfer, learning effect evaluation, and learning correction [26]. The specific design route is demonstrated in Figure 3.

The design of learning objectives and content emphasizes that teachers conduct an in-depth study of course requirements, and teachers are guided to determine teaching objectives. The preassessment analysis focuses on pre-assessing the learner’s level. Preassessment is a prerequisite for inducing DL. It is very important to create a learning cultural atmosphere to stimulate students’ learning interests, so students can actively participate in real learning situations. Learners update and recognize based on their original knowledge in the process of learning. It is necessary to stimulate the students’ prior knowledge and activate the neural network of the brain before knowledge is transferred. Knowledge is processed through DL knowledge and evaluated. Teaching objectives are continuously revised through evaluation feedback.

**3.2. Research Methods.** Expert interviews and questionnaires are combined here. Teachers in charge of sports microcourses are interviewed to clearly understand the current teaching resources of microcourses for physical education majors. Relevant experts in this field are visited to discuss development strategies. Further related research is carried out using the questionnaire survey to understand the actual situation.

**3.3. Questionnaire Design and Distribution.** The subjects of the questionnaire are students of physical education in three



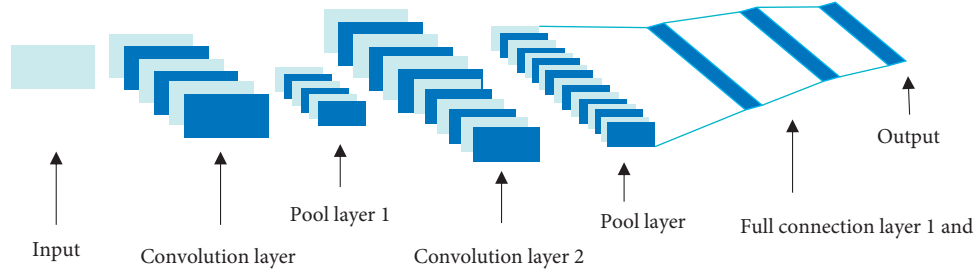


FIGURE 1: DL model.

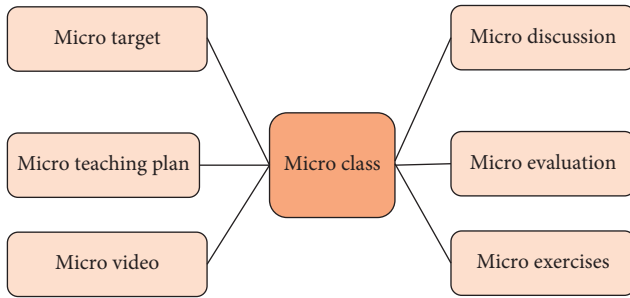


FIGURE 2: Microcourse teaching resources.

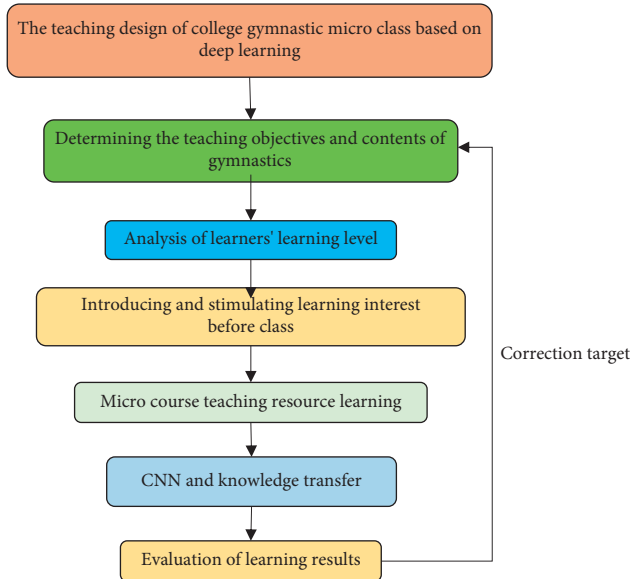


FIGURE 3: Teaching design of college sports gymnastics micro-course based on DL.

different colleges and universities of physical education in Shandong and professional teachers engaged in microcourse teaching of physical education. The three colleges and universities are 985 colleges, 211 colleges, and general colleges. The questionnaire design consists of three parts: the first part is the analysis of the current demand and situation of gymnastics microcourse teaching resources for students majoring in physical education. The second part is a questionnaire for teachers to optimize the teaching resources

of the gymnastics microcourse. The third part is a questionnaire on the application effect of the new teaching model of gymnastics microcourse by students majoring in physical education.

Questionnaires are distributed mainly through the Questionnaire Star platform. The distribution and recovery of questionnaires for students and teachers are listed in Table 1.

The first part of the questionnaire survey is for students, with a recovery rate of 92.31% and an effective rate of 95.83%. The second part of the questionnaire survey is for teachers, the recovery rate is 100%, and the effective rate is 91.67%. The third part investigates the teaching effect, the recovery rate is 96%, and the effective rate is 93.75%.

**3.4. Questionnaire Validity Evaluation.** Ten relevant experts are solicited to evaluate the validity of the three parts of the questionnaire. The results are listed in Table 2.

On the whole, the design of the questionnaire meets the basic survey requirements, and the content of the questionnaire is perfect.

**3.5. Questionnaire Reliability.** A second questionnaire survey is conducted on a part of the same batch of respondents, with an interval of one month. The data are processed using statistical product and service solutions. The correlation is calculated using the Cronbach coefficient equation. The correlation coefficients are  $R^1 = 0.81$ ,  $R^2 = 0.82$ , and  $R^3 = 0.80$  through statistical calculation. The results indicate that the questionnaire meets the statistical test standard and has high reliability. Excel is used to analyze the data. The Cronbach coefficient equation can be expressed as follows:

$$\alpha = \left( \frac{n}{n-1} \right) \left( 1 - \sum \frac{S_i}{S_t} \right). \quad (7)$$

In equation (7),  $\alpha$  is the reliability coefficient and  $n$  is the number of test items.  $S_i$  is the variance of each subject's score for each question and  $S_t$  is the variance of the total score obtained by all subjects.

## 4. Results and Discussion

**4.1. An Analysis of Interest-Oriented Teaching Philosophy.** An interest-oriented survey of current teaching design concepts is conducted on teachers and students respectively. Teachers' attitudes towards teaching concepts and students'

TABLE 1: Questionnaire distribution and recovery.

Classification	The number of questionnaires issued	Number of returned questionnaires	Number of valid questionnaires	Recovery rate (%)	Questionnaire efficiency (%)
Part 1	130	120	115	92.31	95.83
Part 2	24	24	22	100	96.67
Part 3	100	96	90	96	93.75

TABLE 2: Evaluation of questionnaire validity.

Classification	Very perfect	Perfect	Basically perfect	General	Imperfect
Part 1	4	3	2	1	0
Part 2	3	5	1	1	0
Part 3	5	2	2	1	0

interests in different schools are analyzed. The results are shown in Figure 4.

From Figure 4, most teachers prefer an interest-oriented teaching philosophy. The proportion of ordinary colleges and universities decreases slightly, accounting for 89.2%. The highest proportion is 985 colleges. A few teachers think that the current teaching concept is meaningless. Overall, the interest-based teaching philosophy makes sense. In the evaluation of students' interest in learning, the proportion of interested students is relatively small, and the proportion of students who are not interested exceeds 50%. Among the uninterested students, ordinary colleges accounted for 69.9%, 211 colleges accounted for 56.3%, and 985 colleges accounted for 52.6%.<sup>2</sup>

*4.2. Teachers' Self-Evaluation Analysis of Microcourses.* Through the questionnaire, teachers' self-evaluation of sports microlecture is mainly reflected in two aspects, namely, the evaluation of the reasons for the lack of students' interest and the evaluation of the low utilization rate of microlectures teaching resources. Figure 5 shows the specific content of the evaluation.

In Figure 5, A stands for unaccustomed learning style and B stands for lack of initiative. C stands for a single type of homework question and D stands for the lack of supervision and reward mechanism for learning. E means that the quality of the microvideo is not high and F means that the teaching courseware is cumbersome. G means that the microvideo is too long, H means that the discussion area is not active, and I means that the assessment method is single. Figure 5 shows that there are serious problems with the video quality of microlectures. At least 70% of the students responds that there is a problem with the sound quality of the picture, which is also the main reason for disinterest, followed by the lack of supervision and reward mechanisms. Almost 50% of the students are dissatisfied, and the proportion of students who are not used to learning methods is low. The level of student dissatisfaction is greater than that of 211 colleges and 985 colleges. The main reason for the low resource utilization in Figure 5(b) is that the microvideo is

too long, with a proportion exceeding 80%, followed by the inactive discussion area and cumbersome teaching courseware. The last is a single test method. Educational resources need to be further developed, and the quality of teaching should be improved with the goal of learning interest.

## 5. Evaluation of the Application Effect of the DL Teaching Model in the Microcourse Teaching of Physical Gymnastics in Colleges and Universities

*5.1. Comparative Analysis of Students' Interest in Gymnastics Microcourses.* The students' interest in gymnastics microlectures before and after the application of the DL teaching model is analyzed, and the improvement of interest after the application is statistically compared. The results are shown in Figure 6.

From Figure 6(a), the proportion of students who are not interested before the application is large, and the proportion of students who are not interested after the application has decreased significantly. The interest rate increases to 82.4%. After the application, there are still very few students who think it is average, accounting for only 8.1%. From Figure 6(b), the proportion of obvious interest enhancement is more than 60%. The proportion of small interest increases is about 30%. In addition, less than 10% of students think there is no improvement. The effect is most obvious for students in ordinary colleges and universities, with a significant increase ratio of 70.8%, and a proportion of 2.1% who believe that there is no improvement. It is concluded that the design model can greatly improve the learning interest of students in gymnastics microlectures.

*5.2. Designed Model Application Effect Evaluation and Analysis.* The degree of recognition of the students' microvideo learning effect after the model is applied for a year is analyzed. The design model is compared with the teaching methods of MOOC and Chaoxing MOOC in Chinese universities. The results are shown in Figure 7.

From Figure 7, after the application of the design model, the teaching focus has been significantly improved, and the recognition rate has increased from 29% to 82%. Video interest has also been significantly improved, with approval from 25% to 84%. The visual impact and the dynamism of the picture have also been improved to a certain extent. The hearing effect has not improved much, from 41% to 52%. Overall, the design model can effectively improve the learning effect of microvideo. Student recognition has increased significantly. From Figure 7(b), the design model

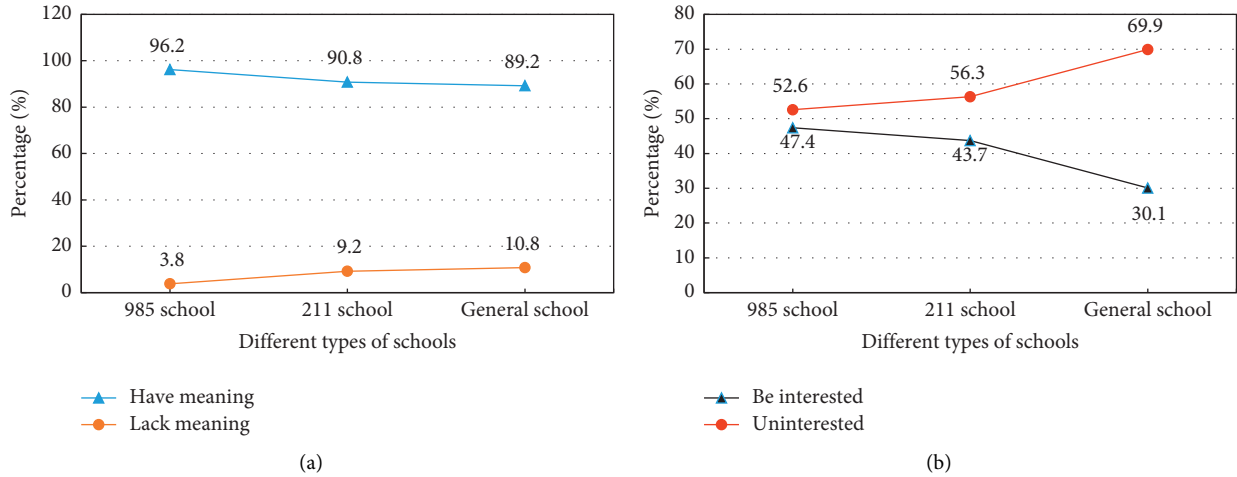


FIGURE 4: Analysis of the current stage of teaching design concept based on interest. (a) Teachers' attitude towards teaching concepts. (b) Students' interest in microcourses.

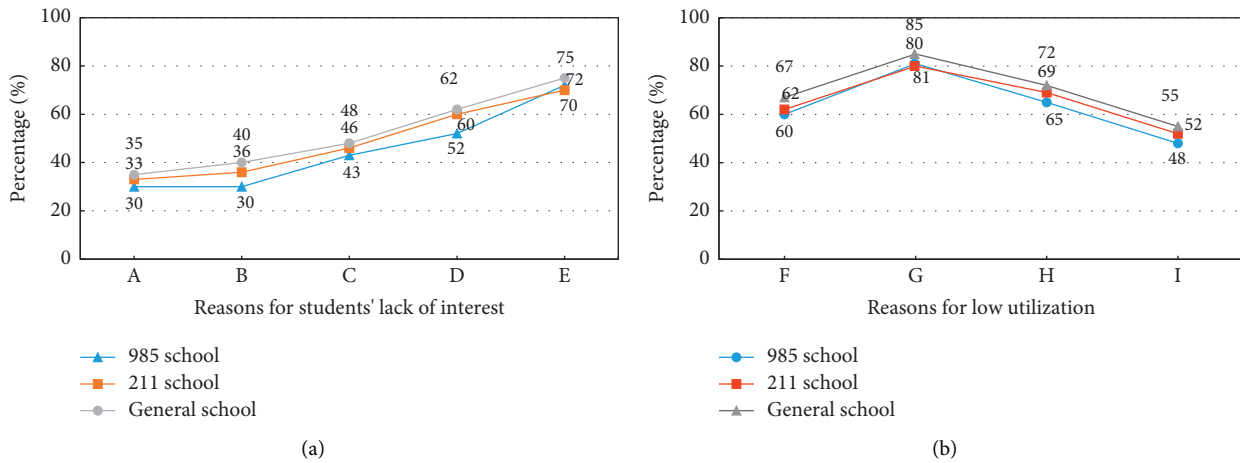


FIGURE 5: Teachers' self-evaluation analysis of microcourses. (a) Evaluation of the reasons for the lack of students' interest. (b) Evaluation of the low utilization rate of microcourse teaching resources.

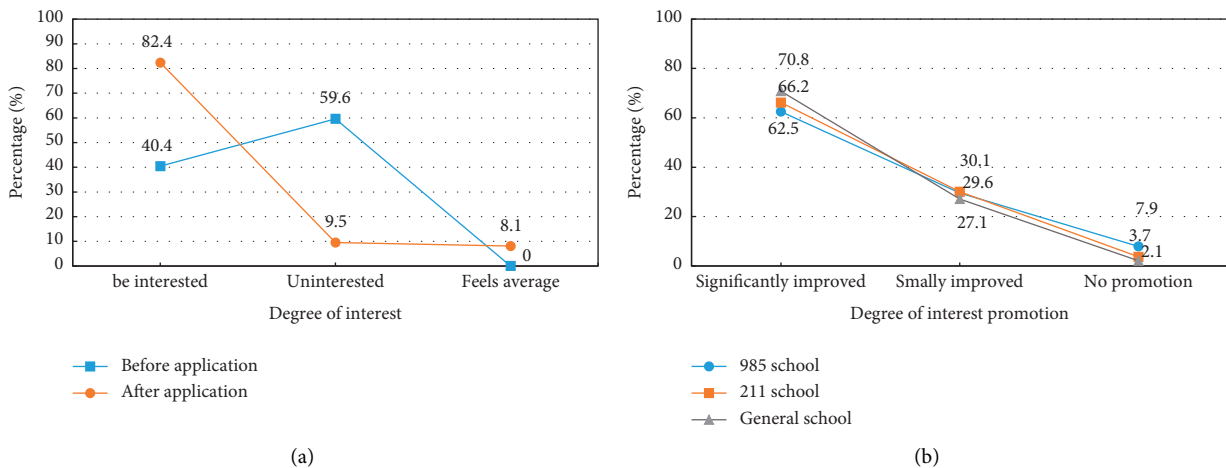


FIGURE 6: Comparison of increased interest in gymnastics microcourses. (a) Analysis of interest degree before and after application. (b) Analysis of interest promotion degree after the application.

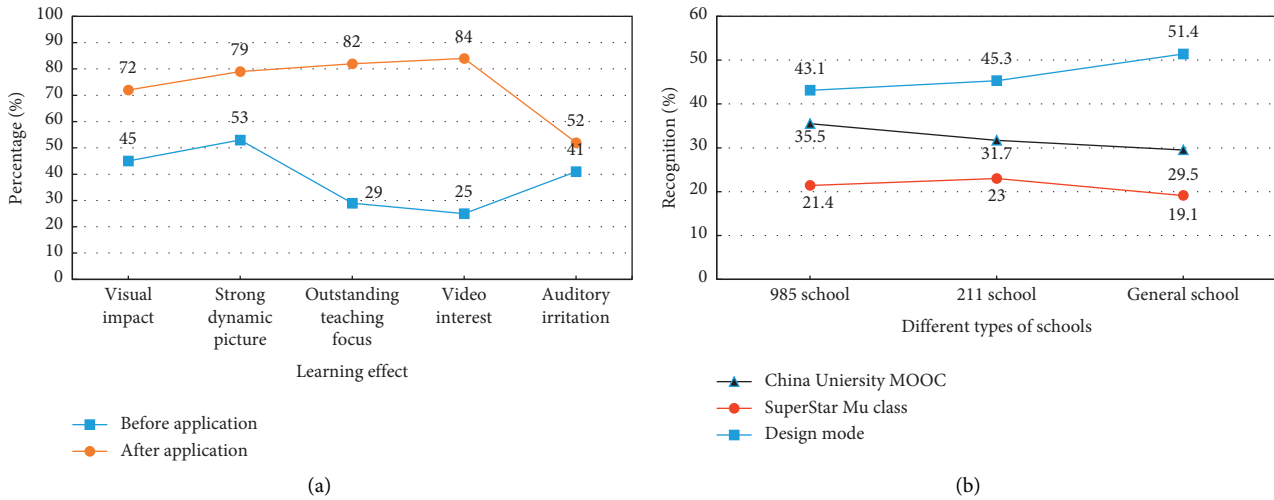


FIGURE 7: Comparative analysis of microvideo learning effects. (a) Microvideo learning effect analysis. (b) Comparative analysis of learning methods.

recognition is higher than that of Chinese university MOOCs and Chaoxing MOOCs, with the highest recognition.

5.3. Analysis of the Help of Designed Models in Learning. The help of the DL design model to students' learning after the application is analyzed, and it is compared with before the application of the designed models. Figure 8 shows the results. Besides, the recognition degrees are compared from five aspects.

After the application, the recognition degree has increased by a certain percentage. The recognition of clear knowledge points has increased from 35% to 42%, and the proportion is not very obvious. Flexibility increases from 21% to 47%, and effective review increases from 24% to 52%. Both have some improvement. The self-learning ability has improved the most, from 18% to 72%. The improvement in learning efficiency has increased from 39% to 82%, a great improvement. Flexibility has also been improved to some extent. As a result, the designed model is helpful to the learning of sports microcourses.

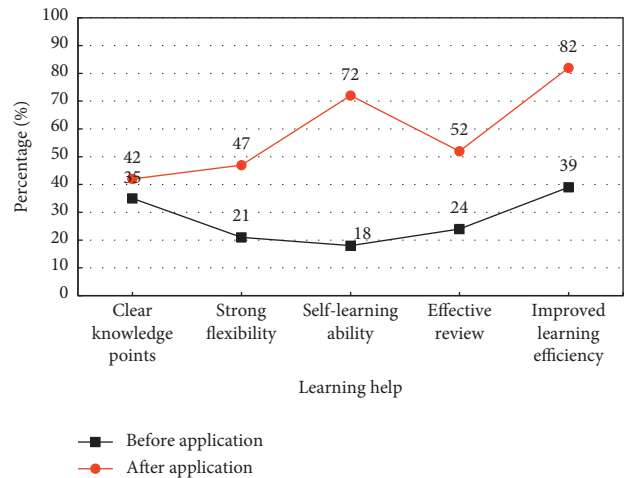


FIGURE 8: Analysis of the help of designed models in learning.

5.4. Satisfaction Analysis of Microcourse Teaching Resources. Students' satisfaction with microcourse teaching resources before and after the application of the design model is discussed through microvideos, microexercises, microdiscussions, and microevaluations. The results are shown in Figure 9.

Figure 9 shows that the designed model can improve students' satisfaction with microvideos, from 62% to 89%. Satisfaction with the microexercises increases from 48% to 77%. Microdiscussion satisfaction increases from 26% to 72%. Microevaluation satisfaction increases from 57% to 82%. Satisfaction with all teaching resources of microcourses has been improved. Students are most satisfied with microvideos followed by microevaluations. The designed model stimulates students' interest in learning and improves

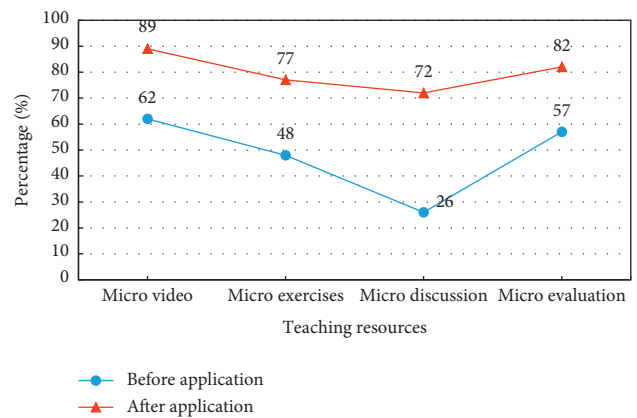


FIGURE 9: Satisfaction analysis of microcourse teaching resources.

the resource utilization of students in gymnastics microcourse teaching. It promotes the teaching and reform of gymnastics courses for sports professionals.

## 6. Conclusion

It is significant to study various teaching methods and design teaching modes with the continuous advancement of the reform of university physical education. This design introduces the FC and studies the DL model. Based on this, the DL theory is combined with the microcourse of college sports gymnastics. The reform of the microcourse teaching of physical gymnastics is achieved through the optimization design of the teaching model. The teaching status and the teaching effect after the application of the design model are evaluated and analyzed through questionnaires. The results are as follows: (1) most teachers prefer an interest-oriented teaching philosophy. In the evaluation of students' interest in learning, the proportion of interested students is relatively small, and the proportion of students who are not interested exceeds 50%. (2) The main reasons for the lack of students' interest are the low quality of microvideos and the lack of supervision and reward mechanisms. The main reasons for the low utilization of teaching resources are too long videos and inactive discussion forums. (3) After the application of the design model, the student's interest in learning has been greatly improved, and the proportion of interest has increased to 82.4%. The improvement effect of ordinary college students is the most obvious, with a proportion of 70.8%. (4) The design model can effectively improve the learning effect of microvideo, and the recognition degree of students is significantly increased. Compared with other learning methods, students have the highest recognition. (5) The designed model has the most obvious improvement in improving learning efficiency and autonomous learning ability. The improvement of self-learning ability increases from 18% to 72%. The learning efficiency improvement increases from 39% to 82%. Satisfaction with all teaching resources of microlectures has been improved. The designed model greatly stimulates students' interest in learning and improves resource utilization. It is suggested that schools should strengthen management and improve the mechanism. Educational technology departments should accelerate the realization of breakthroughs in information technology. This design has reference significance for the construction and optimization of microlectures teaching resources based on technical actions. The disadvantage is that only the learning effect of the design model in the gymnastics microcourse is studied, and it is not involved in other courses of the sports major. The research content is narrow. The next step will be to delve into the application of design models in other sports courses and expand the research object.

## Data Availability

The experimental data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.








## References

- [1] Y. Ding, Y. Li, and L. Cheng, "Application of Internet of Things and virtual reality technology in college physical education," *IEEE Access*, vol. 8, pp. 96065–96074, 2020.
- [2] J. Zhang and H. Mu, "Exploration and practice of flipped classroom in physical education teaching in vocational colleges," *Advances in Educational Technology and Psychology*, vol. 5, no. 11, pp. 67–72, 2021.
- [3] F. J. Hinojo Lucena, J. López Belmonte, A. Fuentes Cabrera, J. M. Trujillo Torres, and S. Pozo Sanchez, "Academic effects of the use of flipped learning in physical education," *International Journal of Environmental Research and Public Health*, vol. 17, no. 1, p. 276, 2019.
- [4] K. Julia, V. R. Peter, and K. Marco, "Educational scalability in MOOCs: analysing instructional designs to find best practices," *Computers & Education*, vol. 161, Article ID 104054, 2021.
- [5] R. Fang, Z. Yang, Y. He, Y. Wang, and H. Zhang, "Effectiveness Evaluation of Physical Education Flipped Classroom Teaching Based on Knowledge Construction," *Mobile Information Systems*, vol. 2022, Article ID 1507167, 12 pages, 2022.
- [6] X. Liao, "The construction and application of "MOOC+ flipped classroom" teaching model for table tennis in colleges and universities," *The International Conference on Cyber Security Intelligence and Analytics*, vol. 114, pp. 576–581, 2020.
- [7] J. Cheng, "Evaluation of physical education teaching based on web embedded system and virtual reality," *Microprocessors and Microsystems*, vol. 83, Article ID 103980, 2021.
- [8] B. Yong, X. Jiang, J. Lin, G. Sun, and Q. Zhou, "Online practical deep learning education," *Educational Technology & Society*, vol. 25, no. 1, pp. 193–204, 2022.
- [9] A. B. Tufail, I. Ullah, W. U. Khan et al., "Diagnosis of diabetic retinopathy through retinal fundus images and 3D convolutional neural networks with limited number of samples," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 6013448, 15 pages, 2021.
- [10] A. B. Tufail, I. Ullah, R. Khan et al., "Recognition of ziziphus lotus through aerial imaging and deep transfer learning approach," *Mobile Information Systems*, vol. 2021, pp. 1–10, Article ID 4310321, 2021.
- [11] R. Khan, Q. Yang, I. Ullah et al., "3D convolutional neural networks based automatic modulation classification in the presence of channel noise," *IET Communications*, vol. 16, no. 5, pp. 497–509, 2022.
- [12] A. B. Tufail, Y. K. Ma, M. K. A. Kaabar et al., "Deep learning in cancer diagnosis and prognosis prediction: a minireview on challenges, recent trends, and future directions," *Computational and Mathematical Methods in Medicine*, vol. 2021, Article ID 9025470, 28 pages, 2021.
- [13] W. Shuo and M. Ming, "Exploring online intelligent teaching method with machine learning and SVM algorithm," *Neural Computing & Applications*, vol. 34, no. 4, pp. 2583–2596, 2022.
- [14] M. A. Jan, X. He, H. Song, and M. Babar, "Editorial: machine learning and big data analytics for IoT-enabled smart cities," *Mobile Networks and Applications*, vol. 26, no. 1, pp. 156–158, 2021.
- [15] K. F. Hew, S. Bai, P. Dawson, and C. K. Lo, "Meta-analyses of flipped classroom studies: a review of methodology," *Educational Research Review*, vol. 33, Article ID 100393, 2021.
- [16] H. Galindo-Dominguez, "Flipped classroom in the educational system," *Educational Technology & Society*, vol. 24, no. 3, pp. 44–60, 2021.

- [17] H. Jasim Mohammed and H. Ati Daham, "Analytic hierarchy process for evaluating flipped classroom learning," *Computers, Materials & Continua*, vol. 66, no. 3, pp. 2229–2239, 2021.
- [18] H. J. Cho, K. Zhao, C. R. Lee, D. Runshe, and C. Krousgrill, "Active learning through flipped classroom in mechanical engineering: improving students' perception of learning and performance," *International Journal of STEM Education*, vol. 8, no. 1, pp. 46–13, 2021.
- [19] Z. Niu, G. Zhong, and H. Yu, "A review on the attention mechanism of deep learning," *Neurocomputing*, vol. 452, pp. 48–62, 2021.
- [20] S. Dong, P. Wang, and K. Abbas, "A survey on deep learning and its applications," *Computer Science Review*, vol. 40, Article ID 100379, 2021.
- [21] P. L. Bartlett, A. Montanari, and A. Rakhlin, "Deep learning: a statistical viewpoint," *Acta Numerica*, vol. 30, pp. 87–201, 2021.
- [22] C. Martins Jarnalo, P. V. M. Linsen, S. P. Blazís, P. van der Valk, and D. Dickerscheid, "Clinical evaluation of a deep-learning-based computer-aided detection system for the detection of pulmonary nodules in a large teaching hospital," *Clinical Radiology*, vol. 76, no. 11, pp. 838–845, 2021.
- [23] D. Ge, X. Wang, and J. Liu, "A teaching quality evaluation model for preschool teachers based on deep learning," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 16, no. 03, pp. 127–143, 2021.
- [24] X. Lei, H. Pan, and X. Huang, "A dilated CNN model for image classification," *IEEE Access*, vol. 7, pp. 124087–124095, 2019.
- [25] A. Onan, "Sentiment analysis on massive open online course evaluations: a text mining and deep learning approach," *Computer Applications in Engineering Education*, vol. 29, no. 3, pp. 572–589, 2021.
- [26] F. J. Hinojo-Lucena, Á. C. Mingorance-Estrada, J. M. Trujillo-Torres, I. Aznar-Diaz, and M. Caceres Reche, "Incidence of the flipped classroom in the physical education students' academic performance in university contexts," *Sustainability*, vol. 10, no. 5, p. 1334, 2018.

## Research Article

# A Natural Language Processing (NLP) Evaluation on COVID-19 Rumour Dataset Using Deep Learning Techniques

**Rubia Fatima** <sup>1</sup>, **Naila Samad Shaikh**,<sup>2</sup> **Adnan Riaz** <sup>3</sup>, **Sadique Ahmad** <sup>4,5</sup>,  
**Mohammed A. El-Affendi** <sup>4</sup>, **Khaled A. Z. Alyamani**,<sup>6</sup> **Muhammad Nabeel**,<sup>7</sup>  
**Javed Ali Khan** <sup>8</sup>, **Affan Yasin** <sup>1</sup> and **Rana M. Amir Latif** <sup>9</sup>

<sup>1</sup>School of Software, Tsinghua University, Beijing, China

<sup>2</sup>Government Degree College for Women, Bosan Road, Multan, Pakistan

<sup>3</sup>Department of Creative Technologies, Faculty of Computing and Artificial Intelligence, Air University, Islamabad, Pakistan

<sup>4</sup>EIAS—Data Science and Blockchain Laboratory, College of Computer and Information Sciences, Prince Sultan University, Riyadh 11586, Saudi Arabia

<sup>5</sup>Department of Computer Sciences, Bahria University Karachi Campus, Karachi, Pakistan

<sup>6</sup>Applied College, Abqaiq Branch, King Faisal University, P.O. Box 4000, Al-Ahsa 31982, Hofuf, Saudi Arabia

<sup>7</sup>School of Software Engineering, South China University of Technology, Guangzhou, China

<sup>8</sup>Department of Software Engineering, University of Science and Technology Bannu, Bannu, Pakistan

<sup>9</sup>Department of Computer Science, COMSATS University Islamabad, Sahiwal Campus, Islamabad, Pakistan

Correspondence should be addressed to Affan Yasin; [affanyasin@tsinghua.edu.cn](mailto:affanyasin@tsinghua.edu.cn)

Received 17 May 2022; Revised 18 June 2022; Accepted 22 July 2022; Published 14 September 2022

Academic Editor: Abdul Rehman Javed

Copyright © 2022 Rubia Fatima et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

*Context and Background:* Since December 2019, the coronavirus (COVID-19) epidemic has sparked considerable alarm among the general community and significantly affected societal attitudes and perceptions. Apart from the disease itself, many people suffer from anxiety and depression due to the disease and the present threat of an outbreak. Due to the fast propagation of the virus and misleading/fake information, the issues of public discourse alter, resulting in significant confusion in certain places. Rumours are unproven facts or stories that propagate and promote sentiments of prejudice, hatred, and fear. *Objective.* The study's objective is to propose a novel solution to detect fake news using state-of-the-art machines and deep learning models. Furthermore, to analyse which models outperformed in detecting the fake news. *Method.* In the research study, we adapted a COVID-19 rumours dataset, which incorporates rumours from news websites and tweets, together with information about the rumours. It is important to analyse data utilizing Natural Language Processing (NLP) and Deep Learning (DL) approaches. Based on the accuracy, precision, recall, and the f1 score, we can assess the effectiveness of the ML and DL algorithms. *Results.* The data adopted from the source (mentioned in the paper) have collected 9200 comments from Google and 34,779 Twitter postings filtered for phrases connected with COVID-19-related fake news. *Experiment 1.* The dataset was assessed using the following three criteria: veracity, stance, and sentiment. In these terms, we have different labels, and we have applied the DL algorithms separately to each term. We have used different models in the experiment such as (i) LSTM and (ii) Temporal Convolution Networks (TCN). The TCN model has more performance on each measurement parameter in the evaluated results. So, we have used the TCN model for the practical implication for better findings. *Experiment 2.* In the second experiment, we have used different state-of-the-art deep learning models and algorithms such as (i) Simple RNN; (ii) LSTM + Word Embedding; (iii) Bidirectional + Word Embedding; (iv) LSTM + CNN-1D; and (v) BERT. Furthermore, we have evaluated the performance of these models on all three datasets, e.g., veracity, stance, and sentiment. Based on our second experimental evaluation, the BERT has a superior performance over the other models compared.

## 1. Introduction

In December 2019, the pandemic for 2019s novel Coronavirus (COVID-19) in Wuhan, China, became a worldwide severe public health problem [1, 2]. “An extreme acute respiratory syndrome called SARS-CoV-2 [3] has often been called the COVID-19 pandemic virus. Coronavirus (CoV) is a broad family of viruses that cause cold diseases, such as Middle East Respiratory Syndrome (MERS-CoV) and SARS-CoV.” The latest strain detected in 2019 and not commonly associated with human illness is COVID-19. Regardless of transmission from humans to wildlife, coronaviruses are zoonotic (a disease that can be transmitted to humans from animals). Studies suggest that a SARS-CoV infection from cats to humans is being transmitted and that the MERS-CoV is tainted with dromedary infection in humans [4]. The virus COVID-19 is believed to be bats-to-human infection. The widespread distribution of the virus culminated in the pulmonary transfer of the disease from person to person. Although, in approximately 82% of cases COVID-19 induces milder effects, some are severe or critical [5]; dyspnoea (shortness of breath), cough, and fever are indicators of infection.

Genome sequencing of respiratory or blood samples may be used to confirm the COVID-19 diagnosis as a major predictor for RT-PCR or in-patient treatment. Although RT-PCR shows poor resistance, many COVID-19 patients remain undetected and unmanageable. Furthermore, the danger of infecting a larger population due to the virus’ high contagiousness cannot be undermined [6]. It is therefore critical that artificial intelligence capabilities be leveraged through the use of ultrasound, X-rays, and computed tomography images through emerging medical care systems that improve diagnostics of COVID-19 [7–9].

Today, diagnosis covers anyone who displays the famous pneumonia trend in COVID-19 chest scan, rather than search for successful results. Our proposed approach would enable policymakers to separate patients and manage them faster. Many people recover with constant lung injury even though there is no mortality with COVID-19. COVID-19 also has the lungs like SARS, which gives them a “honeycomb-like look,” according to the World Health Organization. On the one side, artificial learning leads to product development and, on the other side, to handle global crises. The COVID-19 treatment includes medical instruments and professional personnel who are subject to elevated risks themselves because there is no appropriate managed environment.

Scholars have focused on machine learning NLP methods to prevent the propagation of disinformation [10]. Soni and Roberts [11] identified BERT model with very little preprocessing text, yet achieved excellent efficiency. Facebook deleted more than 50 million posts linked to COVID-19 by April 2020 since they were identified as disinformation using NLP-based machine learning methods. Using these deep learning algorithms [12, 13], social media companies like Twitter and Google have also taken down adverts, and fraudulent posts related to COVID-19 [14].

Although attempts have been carried out utilizing deep learning models to identify COVID-19 disinformation, there

has been a scarcity of research on how ordinary people might simultaneously recognize false information and boost their confidence [15]. Furthermore, black-box models are frequently used in machine-learning-based NLP approaches. Explainable AI in high-risk decision-making is more important in other areas of medicine, such as COVID-19 and fatigue detection. On the other hand, if these models offer insights, they may assist in increasing confidence and acceptance as well as achieving the desired goals [14].

This study utilized several deep learning algorithms, such as LSTM networks, which are redundant neural networks that can learn order dependence in sequence prediction issues. It is necessary for complicated problem areas like machine translation and voice recognition to utilize DL approaches [16, 17].

The key papers initially suggested the video segmentation of the Temporal Convolutional Networks (TCNs). The two stages in this traditional method involve low-level calculating functions using CNN [18–20] to encode spatial-temporal information and introduce low-level functions into a classifier, which captures time information at high levels using (generally) RNN. A similar method needs two distinct models, which is the major drawback [21]. TCN offers a unified method to capture all two information layers (encoder-decoder) hierarchically.

The spread of false information concerning COVID-19 poses a severe risk to public health [22]. Roozenbeek et al. [23] investigate common misconceptions regarding the virus and look into the factors that influence people’s willingness to accept the most widely spread falsehoods. The authors also find that people’s compliance with public health guidelines concerning COVID-19 is negatively affected by their susceptibility to misinformation.

COVID-19 is thus still in significant need of rumours to analyse mood and other rumour categorization activities, including position verification of COVID-19 rumours. We gathered COVID-19 9000+ news rumours and 34,000+ tweets with feelings and labels of position for the research study. Figure 1 illustrates examples of our adapted statistics and data structure. We also analysed our dataset using statistical analysis of rumour propagation and classification findings for a deeper learning rumour classification.

This paper (Figure 2) is divided into five sections. The first part is an introduction to the work that we want to do. The literature review of pertinent research is presented in the second part. The technique and data gathering process are discussed in the third part. The results and discussion of the entire study are presented in the fourth part. We have come to the end of the research in the fifth segment.

## 2. Literature Review

As a result of this study’s usage of COVID-19 from Sina Weibo, it is possible to identify rumours with a smaller number of marked occurrences. The author provides a rumour dataset from Sina Weibo COVID-19 and offers a short, multimodal fusion model to detect rumours. A considerable improvement in rumour identification was



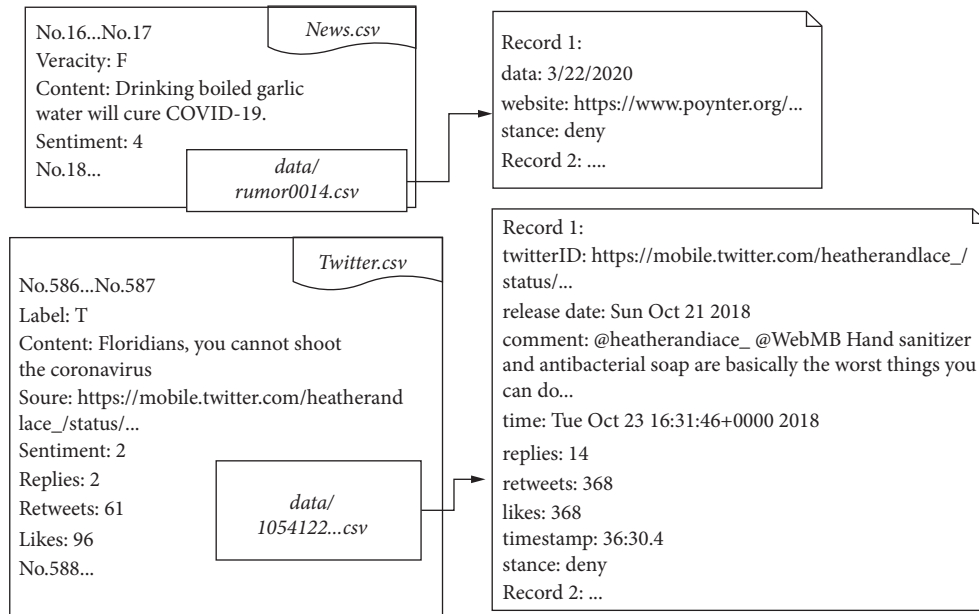


FIGURE 1: Examples of data structure adopted from [24].

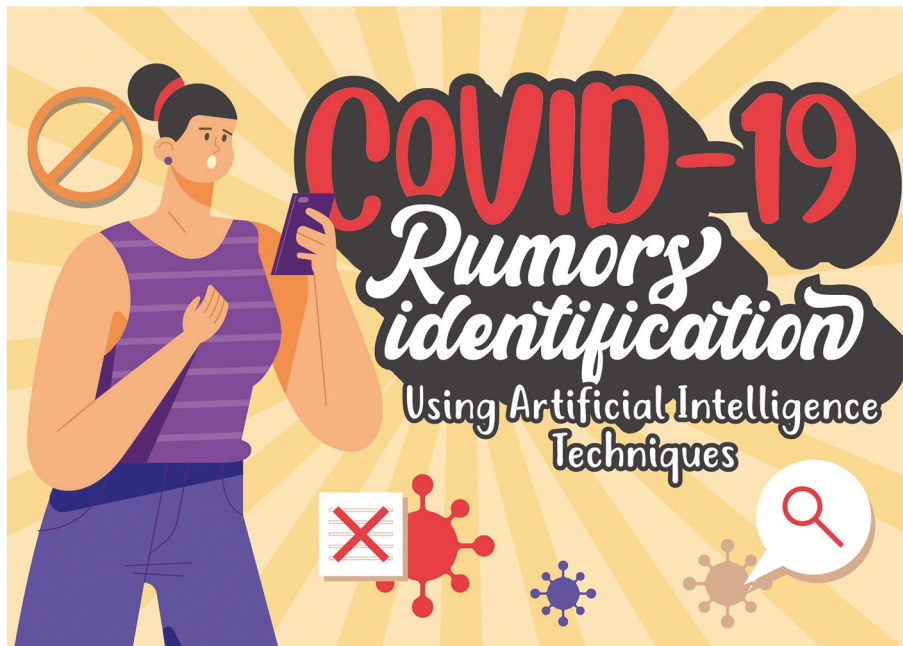


FIGURE 2: COVID-19 rumour identification using AI techniques.

observed on the Weibo and public PHEME datasets acquired by the model [25].

To achieve relatively close news accuracy of a classification and decrease root-mean-square error, the researchers used an actual GitHub dataset framed by COVID-19 news-related parameters. The deep learning classification schema developed the system with the greatest f1 score, which delivers 90% information categorization effectiveness [26, 27]. The author discusses methods of making existing and future approaches to the NLP more inclusive, including

alternate methodologies, using off-the-box technologies, and establishing meaningful collaborations. The author proposes some guidelines for researchers who want to maximize the beneficial social effects of NLP [13, 28].

Authors work on four fundamental tasks of the NLP: retrieval of information, identification of named entities, literature-based discovery, and answering questions. The author also discusses four additional tasks directly addressing elements of the pandemic: topical modelling, sentimental and emotional analysis, predictive caseloads,

and identification of disinformation. Finally, the authors highlight observed trends and difficulties [29].

The epidemic spreads, and more individuals seek COVID-19 testing and therapy. This cybercrime problem will probably persist. Information intelligence can improve the removal and prevention of harmful material to public authorities, regulators, legitimate manufacturers, and technological platforms [30].

The author presents a CORD19STS dataset to resolve this gain, which contains 13,710 sentence pairings that have been taken from the COVID-19 Open Research Dataset (CORD-19) challenge. In particular, the author produced a thousand pairs of sentences using various sample methods. The author utilized a fine-tuned BERT-like language model called Sen SCI-CORD19-BERT to compute similitude values between phrase pairings, which offers us an overall total of 32K phrase pairs to provide a balanced dataset for various semantical similar levels [31].

The evaluation encompasses about 150 NLP research and around 50 COVID-19 datasets and systems. Author's work on four fundamental tasks of the NLP: retrieval of information, identification of named entities, literature-based discovery, and answering questions. The author also discusses four additional tasks directly addressing elements of the pandemic: topical modelling, sentimental and emotional analysis, predictive caseloads, and identification of misinformation [29].

The research [32] reviews many documents that address similar problems with false news, sentiments categorization, and topics extraction. In the article, researchers are directed to valued practices to assist public authorities to fight the increase in falsification and harmful and hatred remarks, which may help enhance present research on COVID-19-related datasets [32].

Like other algorithms in natural language processing, it was suggested that the media articles be categorized as a dataset to evaluate the effects of COVID-19 pandemics in various sectors of the world economy. The model's accuracy was investigated based on the consistency and perplexity score, using LDA algorithms 0.51 and -10.90. Both the algorithm LDA and NMF found common issues in many areas of industry that were affected by the COVID-19 epidemic [33].

This study aims to create a natural language processing pipeline that can recognize patient information based on guidelines, and annotate it with Unify Medical Language Systems ideas for manual physician evaluation. The Human Abstraction, 2513 German clinical notes from the electro-health report, is the gold standard for one-time evaluation. Clinical decision assistance systems might be developed by identifying recommendations from narrative clinical notes [34].

In this research, the author presents an automated summary assessment model (ASE), which is strictly dependent on the characteristics of the source text or the synthesis, which makes a quality model entirely text based. Summaries with accuracy above 80% are successfully classified as low or high quality. The model was created especially on many source texts, which allows for generalization across the text [35].

In 1964, PubMed and EMBASE data search was restricted to 27 suitable items. Data were collected for each research, purpose, the corpus of free texts, patients, symptoms, NLP technique, measurement metrics, and quality indicators. Future NLP research in EHR free-text narratives should study symptoms and symptom documentation. Investigating patient features and publicly developing NLP or pipelines and vocabulary algorithms linked to symptoms [36].

Patients who had chest CTs from 2000 through 2016 were found by interrogating institutional databases at a major quaternary referral centre. Using NLP, imaging reports were identified using GGOs, and further population data were obtained. The NLP examined a broad sample of individuals who had CT chests throughout the research. Provision for age, sex, race, and profession is the demographic characteristic of the GGOs reported [37].

In the Coronaviridae family, COVID-19 (Coron Virus Disease-2019) is a member. No known treatment for an infectious disease wreaks havoc in people's lives and economic and financial institutions throughout the globe. Svc, KNN+NCA, Decision Tree Classifiers, and Naive Bayes Bayesian Classifiers were all surpassed by Random Forests Regressor and Classifier [38].

According to one research, calcium channel blockers were linked to lower in-hospital mortality in patients with COVID-19 infection. The particular discovery was made possible by rapidly tailoring an NLP pipeline to the illness domain. Treatment effects previously undetectable by statistical means were discovered when that information was combined with already structured data [39].

Governments can make better choices if they can correctly predict the number of people infected with this virus. Few hybrid forecasting methods are proposed in this research for the COVID-19 time series. Each model may have different parameters, and Bayesian optimization makes it easier to predict future outcomes. Experiment findings show that deep learning models outperform the benchmark model in short- and long-term predicting scenarios [40, 41].

It would be helpful to have a technology that can correctly identify key COVID-19 clinical ideas from the free language in electronic health data to speed up clinical research. The COVID-19 SignSym was rapidly built using a hybrid method that combined deep-learning-based models with selected lexicons and pattern-based rules. Sixteen healthcare institutions currently use the technology publicly available to researchers as a downloadable package (<https://clamp.uth.edu/covid/nlp.php>) [42].

For finding positively diagnosed COVID-19 patients, VA built a Natural Language Processing pipeline and implemented it to speed up chart reviews. The system's accuracy is assessed at 82.4%, while its recall is calculated at 94.2%. Open-source code for a public-facing implementation has been made accessible to the public. So far, this approach has identified 36.1% of all confirmed positive cases in VA as part of the VA national response [43, 44].

In the last several months, the COVID-19 pandemic lockdown and hashtags were all over social media, with good and negative emotions expressed. Denmark and Sweden, for example, had opposing views on the government decision.

However, unlike their South Asian neighbours, where people exhibited fear and anger, their nation's support was almost universal. The author found a new and innovative method to validate Twitter tweets extracted and analysed using supervised deep learning models [45].

Artificial intelligence has proven its capabilities in Natural Language Processing, Computer Vision, Bio-Informatics, Genetics, and Medical Sciences. The main reasons for such achievements are the availability of high computational resources and the rise of the big data era. However, it requires a deeper comprehension of relevant literature and resources to develop AI systems that help humanity in the real world. Experts in AI have a cross-domain knowledge to solve even simple tasks. Likewise, non-AI professionals have to dig deeper into AI literature to leverage AI capabilities for their domains. Time is a precious resource whose limitation can turn small obstacles into a big disaster, as observed in the COVID-19 epidemic. We believe that widespread infectious cases would be controlled if AI-driven medical-assistive tools could be developed in their early phase. As per our findings, no appropriate open-source platform offers AI services for a medical researcher to come up with exploration, prediction, or classification outcomes to overcome underlying challenges.

### 3. Methodology and Data Collection

**3.1. Method.** The data are taken from the study [24] and <https://github.com/MickeysClubhouse/COVID-19-rumor-dataset>. The summary of the process is explained below for the readers' information and knowledge. The rumours data are gathered from several sources such as Twitter discussions, which featured real-time arguments with specific tags, for data collection. To collect information from Tweets, the author concentrates on a few important hashtags and official accounts (such as NBC, Reuters, CNN, and News Channel) to track the latest developments on hot issues. The author includes the rumour phrases in the database and enhances the dataset with other information, like the website source, the publication date, the validity, the emotion, and the position taken in response to the rumour. The author also records any posts and comments or tweeting of the rumours and the stances taken by the people who shared them. The author divides the gathered rumours into two datasets depending on their source: (i) A bulletin dataset that contains rumours gathered from bulletin sites, and (ii) a tweet dataset that contains rumours gathered from social media platforms such as Twitter.

**3.1.1. Data Collection.** Cheng et al. [24] have created website crawlers that consistently harvest information from the Chrome browser and Twitter. Figure 3 displays the data collection method and the timeline for completion.

**3.1.2. Tweets Gathering Model.** Using COV-19-related hashtags, including COV-19 and coronavirus, Cheng et al. gathered and recorded tweets on COVID-19 in the database. CSV files are available inside fake ID formats, issue date, and

full text. Duplicate tweets were removed from the system. The emotion associated with each rumour is then determined via thoroughly examining the feeling elicited by the rumour's context. Furthermore, the author got the information associated with each tweet, including the text of the reply/retweet remark, the retweet number, the reply number, the like number, and the date the tweet was published. The metadata are then stored in separate files named after the fake IDs used in the tweets.

**3.1.3. News Collecting Method.** Cheng et al. gathered comprehensive information on chosen news items from the Google browser; the authors utilize the mitproxy method that enables HTTPS proxy. It is an interactive HTTPS proxy that is free to use. In contrast to the Tweet crawler, the search engine crawler primarily collects information from the page of search results. It searches the relative and absolute paths of the results based on the URL and timestamp. The author saved the rumours about breaking news in the *news.csv* file; in the *rumourID.csv* file, the author stored the rumours about reposts. Each rumour record includes a truthfulness label and the rumour's substance, and each reposting record includes a repost date, a repost website, and a stance tag, among other things. *It is important to note that not every source online has up-to-date material on the subject.*

**3.1.4. News Dataset Customization.** Cheng et al. [24] combined the news dataset with the COVID-19 rumour dataset. "Very Negative (0), Negative (1), Neutral (2), Positive (3), and Very Positive (4)" are among the five label classes in the dataset. This dataset contains hundreds of news items for various labels. So, based on the classifications, the author scraped a new dataset from Google and combined it with the COVID-19 rumour dataset. The author has now employed five classes in the news dataset, as well as statistical data from the dataset, as shown below:

- (i) "0-1159"
- (ii) "1-1895"
- (iii) "2-1079"
- (iv) "3-3536"
- (v) "4-1531"

**3.1.5. Data Records.** All data are stored in.csv files. For the greatest display, it is recommended that we utilize Utf-8 encoding. Two datasets are created to contain rumours gathered from news and Twitter. Figure 4 lists the tag definitions and provides a collection of information for future reference and research.

### 3.2. Datasets

**3.2.1. News Dataset.** Information from news stories on COVID-19 is included in the news dataset. It includes information on emergencies, public figures' remarks, updates on the coronavirus epidemic, and other information. Each

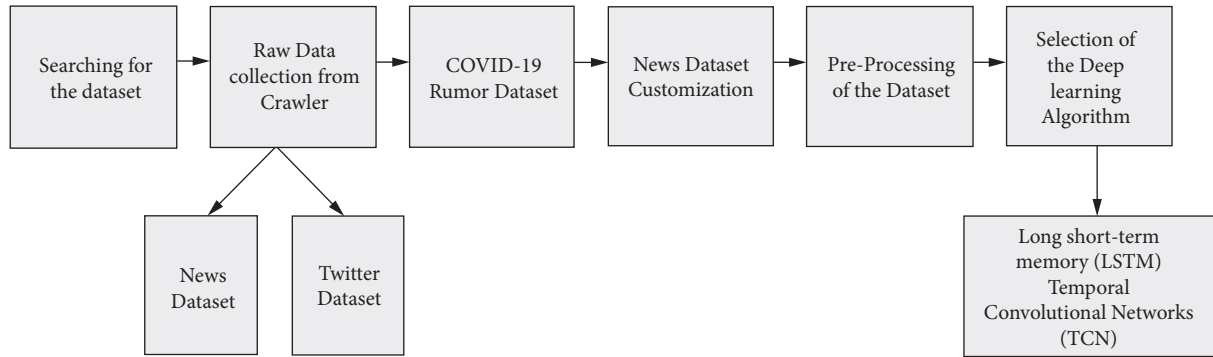


FIGURE 3: Flowchart depicting the collecting, labelling, and postprocessing of datasets.

Term	Label	Explanation and examples
Veracity	True (T)	The content is logical and describing the facts, e.g., “Wuhan has been quarantined.”
	False (F)	The content is made up, or contains false information, e.g., “Drinking bleach can cure coronavirus.”
	Unverified (U)	The authenticity or truthfulness of the statement is hard to judge at the time of labeling.
Stance	Support	Positive attitudes about the content, e.g., “I think the statement is right.”
	Deny	Denying attitudes about the content, e.g., “Are you kidding? This is wrong!”
	Comment	No obvious stance, e.g., “This message is interesting.”
	Query	Doubting the validity of news/tweets, e.g., “Is that ture?” or “Can you prove?”
Sentiment	Very Negative (0)	The content has a strong pessimism.
	Negative (1)	The emotion is pessimistic but weaker than “very negative”
	Neutral (2)	The comment/report is in a plain and narrative tone.
	Positive (3)	The comment reflects positive emotions of aims, such as news providing tips to fight the virus.
	Very Positive (4)	Cheerful news such as progress in the research, massive donations of breakthroughs in the vaccine.

FIGURE 4: Labels in the dataset are highlighted—adopted from [24].

record includes the following structured metadata, which describes the specifics of the news story:

- (i) *Sources*: Some websites include the rumour phrase. The rumour’s origin is determined by counting the number of websites that discuss it, such as those that debate its validity. The oldest rumour source is indicated as the source of the rumour.
- (ii) *Popularity*: The number of websites that repost the entire rumour in the Chrome browser shows the rumour’s popularity.
- (iii) *Date*: The date on which each rumour record was published, as determined by the web crawler, is mechanically gathered.
- (iv) *Stance*: The mindset of the rumour source’s author or editor is important to note. Our categorization system follows the traditional classification system and divides rumour attitude into four categories: support, denial, remark, and inquiry. The positions are labelled and cross-validated manually, after which they are checked against the context of each page. It is worth mentioning that the majority of the positions fall into the support and comment categories.
- (v) *Sentiment*: To identify a rumour term, we need one of five distinct sentiments: very (negative), negative, (neutral), (positive), or (very positive). The author

Cheng et al. carefully classify and cross-validate the feeling to decide whether this is bad or good news, depending on the circumstances. News articles that report: (i) COVID-19 cases are categorized as Negative news; (ii) COVID-19-related deaths are categorized as very negative news; (iii) COVID-19-related prevention tips are categorized as positive news; (iv) COVID-19-related campaigns and vaccine advancement are categorized as highly positive news.

- (vi) *Veracity*: True or false, indicating that the data are describing a reality; unverified, indicating that the news has not been verified as of the time of collection; or true or false, indicating that the news has not been confirmed as of the time of collection. The labelling and cross-validation are done manually during the data collection stage, using reputable sources and widely shared common knowledge [24].

**3.2.2. Twitter Data.** The Twitter dataset includes speculations that have been published on the social media platform. The data are compiled from public accounts which have commented on COVID-19-related information discussion forums that have been labelled with COVID-19-related tags. Other users on the social media site may retweet or reference the discussions.

- (i) In the same way, as sources in a news dataset shows how individuals react to a tweets (from social media twitter), responses indicate what people react to the tweet.
- (ii) Reply/Retweet/Like (RRL) number: These figures show the trend in the spread of a tweet. The crawler automatically parses the RRL number and displays it.
- (iii) Popularity: When you add up the RRL numbers, you get an idea of how popular a tweet is.
- (iv) Data: In the Twitter platform, the date is expressed using MM.DD.YYYY, representing when the tweet was published on the platform [24].

### 3.3. Preprocessing Steps

- (1) Read the dataset
- (2) Removal of stop words
- (3) Removal of symbols
- (4) Removal of digits
- (5) Tokenization from each row
- (6) Making a vector shape
- (7) Pass vector from module

## 4. Results and Discussion

### 4.1. Experiment 1

4.1.1. *Long Short-Term Memory (LSTM)*. LSTM is an in-depth learning artificial recurrent neural network architecture (RNN). In contrast to conventional neural networks, feedback connections are available to LSTM. It can handle single data points, like pictures, and whole data sequences, like voice or video [46]. For example, LSTM is for tasks like unsegmented handwriting recognition, voice recognition, and network traffic or IDS anomaly detection, as shown in Figure 5.

4.1.2. *Temporal Convolutional Networks (TCN)*. The first major study suggested a TCN for the segmentation of video-based actions. The two stages of this classic approach include calculating low-level CNN characteristics that encode space-time information and introducing those low-level characteristics into a ranking that collects time information on a high-level basis using RNN [20, 48, 49]. A similar method needs two distinct models, which is the major drawback. TCN offers a unified method to hierarchically capture all two information layers (encoder-decoder), as shown in Figure 6.

We discussed the analysed findings from several deep learning algorithms in the results section, which we applied to the “news from Google and Twitter” dataset. We employed Keras innovative models “Long short-term memory (LSTM) and Temporal Convolutional Networks (TCN)” in this study. We gathered 9200 Google comments and 34,779 Twitter postings filtered for phrases connected with bogus news about the COVID-19. The dataset was assessed using three criteria: truthfulness, stance, and sentiment. The analysis results reveal which model has the best

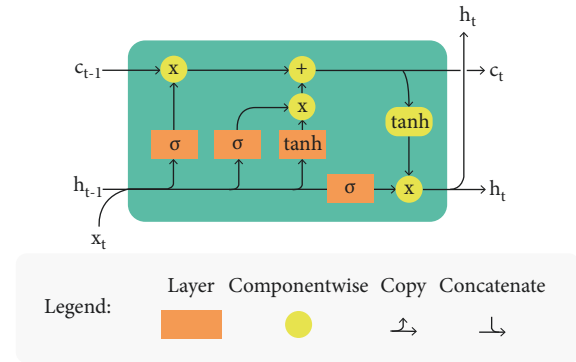


FIGURE 5: Schematic of the long short-term memory cell—adopted from [47].

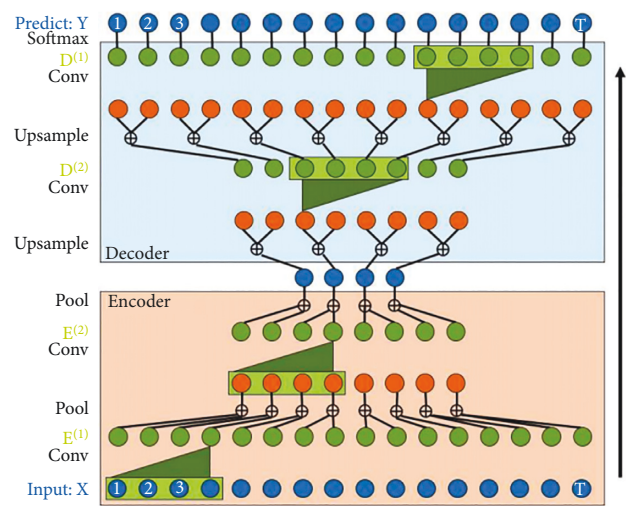


FIGURE 6: Schematic of the Temporal Convolutional Networks—adopted from [50].

accuracy and loss rate. Precision, f1 score, and recall are the additional characteristics. We utilized it to verify the deep learning model’s legitimacy.

4.1.3. *The Evaluated Results for Sentiment and Veracity*. We analysed the dataset of Google rumour news on the COVID-19 in this area. This dataset was used in two ways: sentimentality analysis and truthfulness analysis. “Very Negative (0), Negative (1), Neutral (2), Positive (3), and Very Positive (4)” are the sentiment labels we utilized. “True (T), False (F), and Unverified (U)” is also the veracity. For the analysis, we used 9200 rumour news after preprocessing against COVID-19. A total of two methods for deep learning were used in this study: “Long short-term memory (LSTM)” and “Temporal Convolutional Networks (TCN).” We compared them and determined which one is the best for real-time implementation.

(1) *Compiling Model*. The initial stage is to give the input dataset to several models. The computational graph represents the dropout and layers involved in the model. These models have been trained and tested, and their performance

has been quantified in terms of loss and accuracy. The compared models are implemented using the following hyperparameter settings.

- (i) Batch Size: 5
- (ii) Epoch: 10
- (iii) Optimizer: “Adam”

Each of the models examined in this study is represented in Table 1 by its accuracy and loss rate, with the TCN model having the best val\_loss and val\_accuracy of 0.7345 and 63.59%, respectively. We showed that the TCN model had the best evaluated results for the sentiment dataset since other models’ accuracy and loss rate were included in the table. For each sentiment model, accuracy and loss are shown graphically in Figures 7 and 8.

Table 2 shows the accuracy and loss rate of each model we employed in the investigation, with the TCN model having the best val\_loss and val\_accuracy of 1.3806 and 75.43%, respectively. According to the veracity dataset’s accuracy and loss rate of other models, we found that the TCN model had the best evaluated results. Figures 9 and 10 depict a graphical depiction of the accuracy and loss of all of the veracity models we have developed.

Consequently, our dataset is unbalanced, with 346 of the 920 instances falling into one of the three categories. For each approach, the ratio is 0.73%. Since the predictor is mostly right when it comes to class 3 samples, it has an incredibly high level of precision, recall, and f1-score values for class 3 and extremely low scores for the other classes. Macro  $F1$ ’s objective is to calculate the  $F1$  split by class without using weights for the aggregate:

$$F_{\text{class1}} + F_{\text{class2}} + F_{\text{class3}} + F_{\text{class4}}. \quad (1)$$

You will be penalized if the model fails to perform effectively among minority groups (which you want when there is an imbalance).

It is calculated by the number of true labels in each class when  $F1$  costs are merged:

$$F_{\text{class1}} * C1 + F_{\text{class2}} * C2 + F_{\text{class3}} * C3 + F_{\text{class4}} * C4. \quad (2)$$

Consequently, you prefer the class labels (which you typically would not want).

Therefore, our modelling is incorrect for one of the classes since the macro  $F1$  score correctly captures but is not weighted, leading to the five gaps in our dataset, as shown in Table 3. The TCN model has high precision, recall,  $F1$  score from other models, and accuracy of 64%.

As a result, our dataset is skewed, with 456 out of 920 instances falling into the  $T$  group (0.81% for various techniques, respectively). These results in exceptionally high precision, recall, and f1-score values for class  $T$ , and extremely low scores for the other classes, as a consequence of the predictor virtually always accurately predicting any given sample from class  $T$ . Still, weighted does not result in a mismatch between the three classes in your model. Based on these data, Table 4 shows that the TCN model has a high accuracy of 75% compared to other models.

*4.1.4. The Evaluated Results for Stance.* We analysed the dataset of Google rumour news on the COVID-19 in this area. This dataset was used in two ways: sentiment and truthfulness. “Deny, Comment, Query, and Support” are the sentiment labels we have utilized. We used the 34,779 twitter comments against the COVID-19 after preprocessing for the analysis. For the investigation, we employed two deep learning algorithms: “Long short-term memory (LSTM)” and “Temporal Convolutional Networks (TCN).” We compared them and determined which one is the best for real-time implementation.

The accuracy and loss rate of each model we utilized for the analysis are shown in Table 5. We found that the LSTM model has 64.20% greater accuracy than the TCN model but has more val\_loss, consistent with epoch 4. As a result, we determined that the TCN model had the best val\_loss and val accuracy, with values of 0.6985 and 48.96%, respectively. We examined that the TCN model has the best-evaluated results for the stance dataset since other models’ accuracy and loss rates are shown in Table 5. Figures 11 and 12 depict a graphical depiction of the accuracy and loss of all the deployed stance models.

Consequently, our dataset is unbalanced, with the Deny class accounting for 2462 of the 3478 cases (0.77 and 0.64% for various methods, respectively). Because of this, the class Deny has extraordinarily high accuracy, recall, and  $f1$  scores, while the other classes have extremely poor accuracy, recall, and  $f1$  scores. You can see this in your macro  $F1$  score (correct), but weighted (inaccurate) shows the four-class disparity. Table 6 demonstrates that the TCN model has good accuracy, recall, and  $F1$  score compared to other models and a 49% success rate.

*4.2. Experiment 2.* Table 7 illustrates the sentiment dataset which has been divided into five classes and each class is assigned a specific numerical label value ranging from 0 to 4 describing the extent of how positive or negative the news is. All the classes show varying dataset lengths among which class 2 depicts the lowest length (1079) while class 4 depicts the highest length (1531). The description of all the labels is as follows:

- (i) 0 = very negative
- (ii) 1 = negative
- (iii) 2 = neutral
- (iv) 3 = positive
- (v) 4 = very positive

Table 8 illustrates the stance dataset which is divided into four classes under the labels of comment, support, query and denies. The labels showed significant variations in lengths. The Comment label showed the highest length (24222) while the Deny label exhibited the lowest length (1750).

Table 9 demonstrates the veracity dataset which is divided into four classes including true ( $T$ ), false ( $F$ ), unverified ( $U$ ), and Twitter ( $U$ ). The highest length was shown by  $T$  class (4485) while the lowest length was depicted by Twitter ( $U$ ) (1).

TABLE 1: The evaluating results of the sentiment dataset.

Long short-term memory (LSTM)			Temporal Convolutional Networks (TCN)		
Epoch	Val_Loss	Val_Accuracy	Epoch	Val_Loss	Val_Accuracy
1	0.9743	0.5598	1	0.3340	0.5598
2	0.9648	0.6043	2	0.3490	0.6152
3	1.1173	0.6196	3	0.3237	0.6359
4	1.1919	0.5967	4	0.3680	0.6478
5	1.5083	0.6022	5	0.4400	0.6348
6	1.7218	0.6239	6	0.4447	0.6337
7	1.6471	0.6000	7	0.4657	0.6065
8	1.8342	0.6250	8	0.6769	0.6196
9	1.8323	0.6174	9	0.7163	0.6087
10	1.8093	0.6293	10	0.7345	0.6359

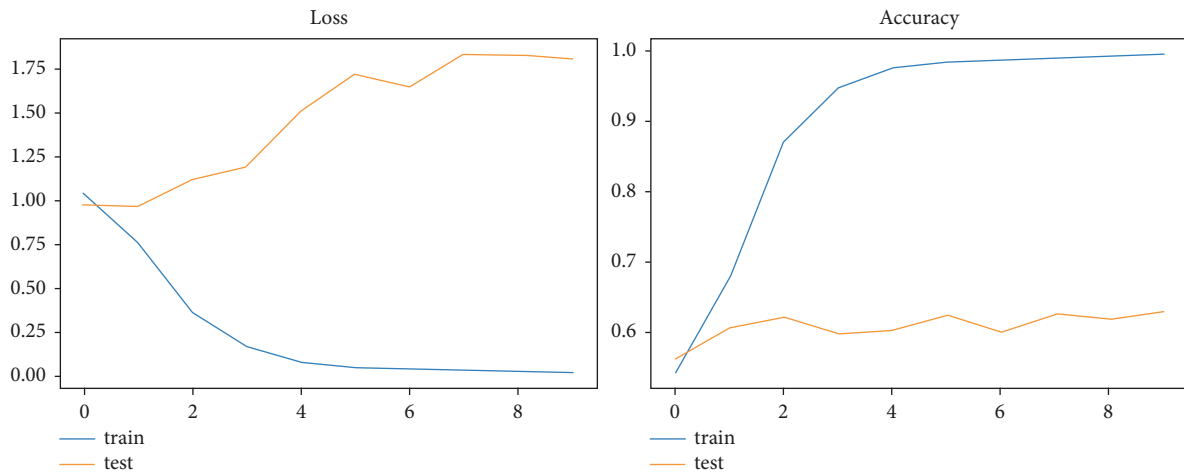


FIGURE 7: LSTM—the graphical representation of evaluating results for sentiment dataset.

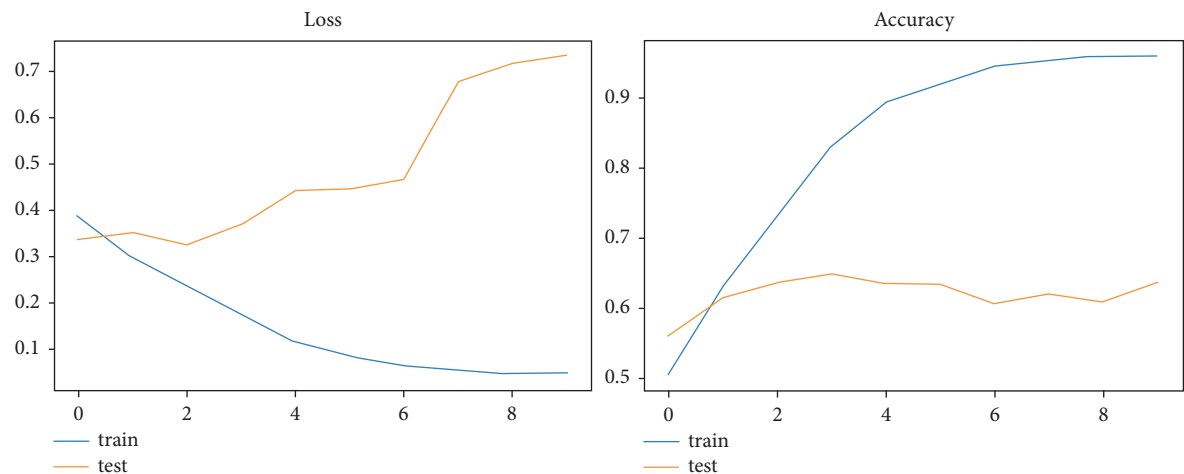


FIGURE 8: TCN—the graphical representation of evaluating results for sentiment dataset.

The datasets presented above showed multiclass division. Therefore, the current study also employed multiclassification models for the evaluation of the datasets. The current study employed several state-of-the-art deep learning models for this purpose.

4.2.1. Training Architecture of Deep Learning Models.

The study used the following deep learning models:

- (i) Simple RNN Architecture
- (ii) LSTM + Word Embedding (WE)

TABLE 2: The evaluating results of the veracity dataset.

Long short-term memory (LSTM)			Temporal Convolutional Networks (TCN)		
Epoch	Val_Loss	Val_Accuracy	Epoch	Val_Loss	Val_Accuracy
1	0.5035	0.7815	1	0.3569	0.7717
2	0.5381	0.7804	2	0.3208	0.7641
3	0.6837	0.7696	3	0.4927	0.7620
4	0.9285	0.7837	4	0.7023	0.7522
5	0.9540	0.7793	5	0.8169	0.7511
6	1.0478	0.7750	6	0.7995	0.7511
7	1.1807	0.7522	7	0.7916	0.7598
8	1.4689	0.7435	8	2.2497	0.7587
9	1.4431	0.7435	9	1.9047	0.7598
10	1.2965	0.7326	10	1.3806	0.7543

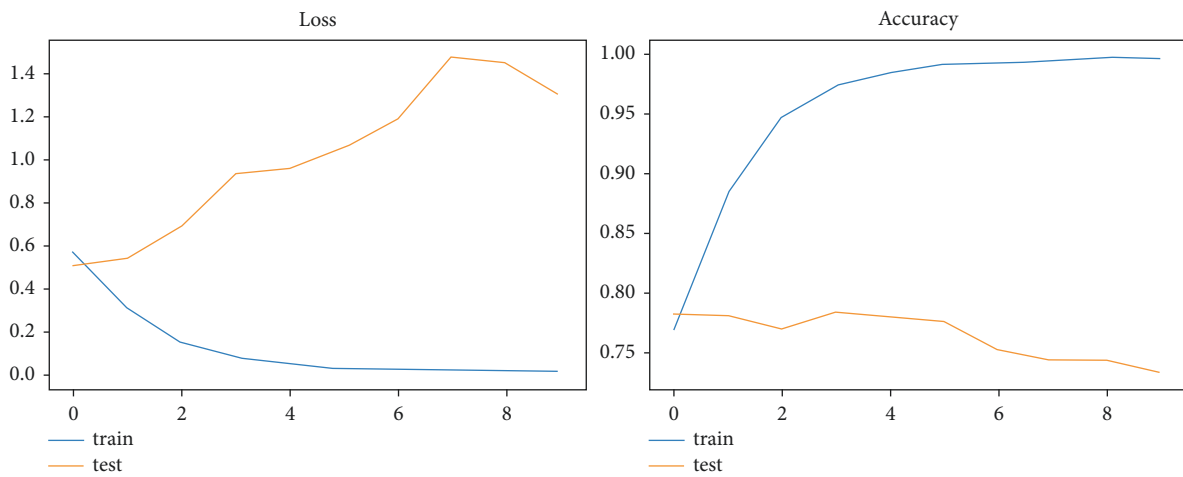


FIGURE 9: LSTM—the graphical representation of evaluating results for veracity dataset.

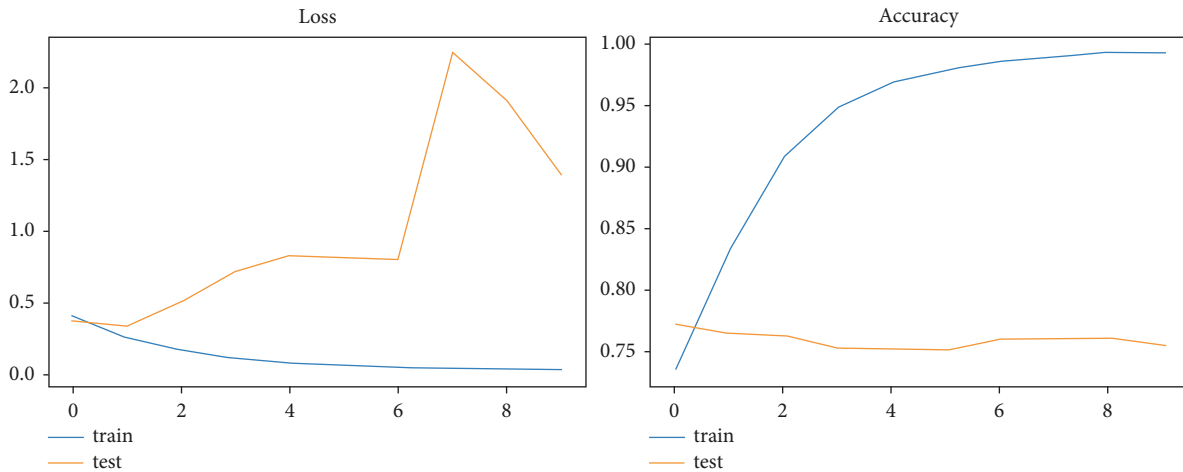


FIGURE 10: TCN—the graphical representation of evaluating results for veracity dataset.

(iii) Bidirectional + Word Embedding (WE)

(iv) LSTM + CNN-1D

The architecture of the important deep learning models is presented below:

(1) *Simple RNN Architecture.* It is a form of neural network in which the nodes are connected with each other. This connection exists along a temporal sequence which refers to the transition of data along with time. This neural network is preferred over conventional neural networks. The reason lies



TABLE 3: The average evaluating results of sentiment dataset.

<!--Col Count:9	Long short-term memory (LSTM)				Temporal Convolutional Networks (TCN)			
	Precision	Recall	F1 score	Support	Precision	Recall	F1 score	Support
0	0.52	0.39	0.45	112	0.56	0.48	0.52	112
1	0.65	0.69	0.67	197	0.62	0.70	0.66	197
2	0.47	0.46	0.47	117	0.44	0.47	0.45	117
3	0.69	0.78	0.73	346	0.71	0.75	0.73	346
4	0.62	0.52	0.56	153	0.71	0.54	0.61	153
Accuracy			0.63	920			0.64	920
Macro avg	0.59	0.57	0.58	920	0.61	0.59	0.59	920
Weighted avg	0.62	0.63	0.62	920	0.64	0.64	0.63	920

TABLE 4: The average evaluating results of the veracity dataset.

	Long short-term memory (LSTM)				Temporal Convolutional Networks (TCN)			
	Precision	Recall	F1 score	Support	Precision	Recall	F1 score	Support
<i>F</i>	0.85	0.74	0.79	343	0.80	0.84	0.82	343
<i>T</i>	0.79	0.82	0.81	456	0.85	0.77	0.81	456
<i>U</i>	0.32	0.40	0.36	121	0.37	0.45	0.41	121
Accuracy			0.73	920			0.75	920
Macro avg	0.65	0.65	0.65	920	0.67	0.69	0.68	920
Weighted avg	0.75	0.73	0.74	920	0.77	0.75	0.76	920

TABLE 5: The evaluating results of the stance dataset.

Long short-term memory (LSTM)			Temporal Convolutional Networks (TCN)		
Epoch	Val_Loss	Val_Accuracy	Epoch	Val_Loss	Val_Accuracy
1	0.7194	0.7432	1	0.3461	0.7430
2	0.7765	0.7303	2	0.3302	0.7418
3	0.9348	0.6691	3	0.3602	0.7266
4	1.0246	0.6722	4	0.4021	0.7033
5	1.3052	0.5828	5	0.4575	0.6455
6	1.3175	0.5828	6	0.5456	0.5808
7	1.4237	0.6650	7	0.5838	0.5555
8	1.5037	0.6676	8	0.7077	0.5664
9	1.6065	0.6834	9	0.6613	0.5428
10	1.7549	0.6420	10	0.6985	0.4896

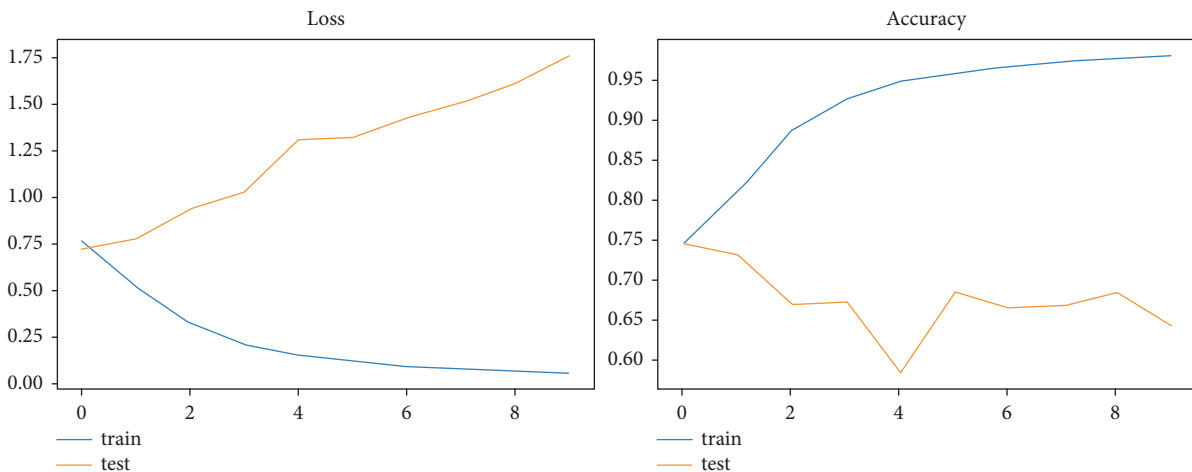


FIGURE 11: LSTM—the graphical representation of evaluating results for stance dataset.

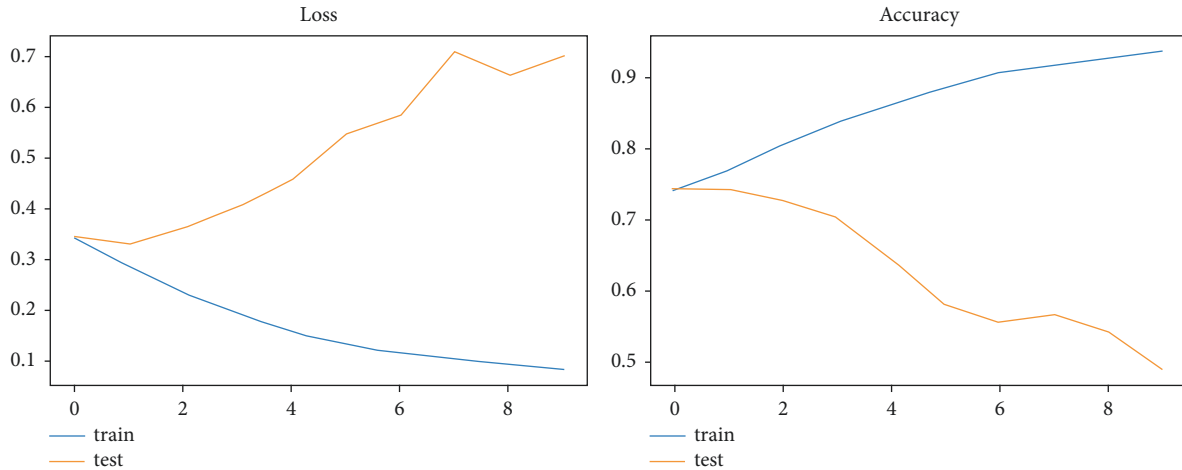


FIGURE 12: TCN—the graphical representation of evaluating results for stance dataset.

TABLE 6: The average evaluating results of stance dataset.

	Long short-term memory (LSTM)				Temporal Convolutional Networks (TCN)			
	Precision	Recall	F1 score	Support	Precision	Recall	F1 score	Support
Deny	0.78	0.76	0.77	2462	0.78	0.54	0.64	2462
Comment	0.32	0.30	0.31	193	0.41	0.09	0.15	193
Query	0.19	0.19	0.19	261	0.08	0.40	0.14	261
Support	0.39	0.47	0.43	562	0.53	0.43	0.48	562
Accuracy			0.64	3478			0.49	3478
Macro avg	0.42	0.43	0.42	3478	0.65	0.65	0.35	3478
Weighted avg	0.65	0.64	0.65	3478	0.75	0.73	0.55	3478

TABLE 7: Statistics of sentiment in the COVID-19 dataset.

Classes	Length
0	1158
1	1895
2	1079
3	3536
4	1531

TABLE 8: Statistics of stance in the COVID-19 dataset.

Classes	Length
Comment	24222
Support	5248
Query	2474
Deny	1750

TABLE 9: Statistics of veracity in the COVID-19 dataset.

Classes	Length
<i>F</i>	3460
<i>T</i>	4485
<i>U</i>	1253
<i>U</i> (twitter)	1

in the capability of this model to process the past input to address the future inputs. For instance, RNN models have the capacity to predict the next word by analysing the

previous words present in the sequence while a conventional model lacks this capability. The RNN models recur the information in a loop which makes it possible for the information to remain in the system. Figure 13 demonstrates the detailed summary of the simple recurrent neural network (RNN) model architecture. The first layer is the embedding layer for which the calculated parameters are 5701200. The next layer is the simple RNN layer for which we calculated 40100 trainable parameters. The additional 505 parameters added in the total param count were due to the dense layer added after the simple RNN layer.

(2) *LSTM + Word Embeddings Architecture*. Long short-term memory (LSTM), a type of recurrent model, is a widely used model in deep learning. This model is highly advantageous because it utilizes both the previous and future memory of the available data. Moreover, its performance in the case of time-series data is also quite effective. We have used Global Vectors for Word Representation (GloVe) for word embeddings which is an unsupervised learning algorithm. These are Google-made word embeddings with almost 800 billion words and 300-dimensional embeddings of them (see Figure 3). Figure 14 shows the detailed model summary of the LSTM + Word Embeddings model used on the dataset of this study. It has three layers including embedding, LSTM, and dense layers. The embedding layer has the same number of parameters as in the previous model (5701200) but they are treated as nontrainable parameters in this model while the dense layer has 505 trainable parameters. However, this model has a different count of the total

Model: "sequential\_6"

Layer (type)	Output Shape	Param #
embedding_6 (Embedding)	(None, 200, 300)	5701200
simple_rnn_6 (SimpleRNN)	(None, 100)	40100
dense_6 (Dense)	(None, 5)	505

=====  
 Total params: 5,741,805  
 Trainable params: 5,741,805  
 Non-trainable params: 0

FIGURE 13: Summary of simple RNN model.

Model: "sequential\_14"

Layer (type)	Output Shape	Param #
embedding_14 (Embedding)	(None, 200, 300)	5701200
lstm_4 (LSTM)	(None, 100)	160400
dense_14 (Dense)	(None, 5)	505

=====  
 Total params: 5,862,105  
 Trainable params: 160,905  
 Non-trainable params: 5,701,200

FIGURE 14: Summary of LSTM + Word Embedding model.

Model: "sequential\_6"

Layer (type)	Output Shape	Param #
embedding_6 (Embedding)	(None, 200, 300)	5652000
bidirectional (Bidirectional)	(None, 600)	1442400
dense_6 (Dense)	(None, 3)	1803

=====  
 Total params: 7,096,203  
 Trainable params: 1,44,203  
 Non-trainable params: 5,652,000

FIGURE 15: Summary of Bidirectional + Word Embedding model.

number of parameters than the previous model because the LSTM layer has 160400 parameters that increase the total number of trainable parameters to 5862105.

(3) *Bidirectional + Word Embeddings Architecture.* We employed the bidirectional feature in combination with word embeddings to function on both the right and left sides of the text dataset. This kind of model works efficiently in cases where the text dataset is quite large and the purpose is to create a summary of the dataset. Figure 15 illustrates the summary of the Bidirectional + Word Embedding model. The model entails three layers, i.e., embedding, bidirectional, and dense layers. The first layer is the embedding layer that has 5652000 nontrainable parameters. The next layer is the

bidirectional layer that has 1442400 trainable parameters. The third layer (dense layer) provides additional 1803 parameters, taking the total count of parameters to 7096203.

(4) *LSTM + CNN-1D Architecture.* It is important to understand the use of both CNN and LSTM to comprehend their combined model use in this study. CNNs have been extensively utilized in modelling issues in relation to inputs, such as image datasets. Over the years, CNNs have provided a great opportunity to detect and classify the image dataset to extract vital information from the dataset. On the other hand, LSTMs are employed in tasks where the dataset has a sequence and they perform predictions based on the sequence. It assists in undertaking those tasks that require image sequences to predict

Model: "sequential\_5"

Layer (type)	Output Shape	Param #
embedding_5 (Embedding)	(None, 200, 128)	2560000
conv1d_10 (Conv1D)	(None, 200, 128)	49280
max_pooling1d_10 (MaxPooling)	(None, 50, 128)	0
conv1d_11 (Conv1D)	(None, 50, 64)	24640
max_pooling1d_11 (MaxPooling)	(None, 24, 64)	0
lstm_5 (LSTM)	(None, 64)	33024
dropout_5 (Dropout)	(None, 64)	0
dense_10 (Dense)	(None, 64)	4160
dense_11 (Dense)	(None, 5)	325
Total params: 2,671,429		
Trainable params: 2,671,429		
Non-trainable params: 0		

None

FIGURE 16: Summary of LSTM + CNN-1D model.

certain information for which a more sophisticated model is required. To this end, the LSTM + CNN model is utilized to predict spatial input such as images and videos. In this model, the task of feature extraction is undertaken using CNN while LSTM helps in prediction. Since the current model is applied to the text dataset, simply a one-dimensional (1D) model is employed. Figure 16 demonstrates the model summary of LSTM + CNN 1D architecture. It contains six types of layers including embedding, Conv1D, max pooling, LSTM, dropout, and dense layers. The model has no nontrainable parameters. The first layer is the embedding layer with the highest number of trainable parameters (2560000) followed by the second layer Conv1D. The LSTM layer contains only 33024 trainable parameters while the max pooling has no trainable parameters that impact backpropagation.

**4.3. Accuracy Performances of Deep Learning Models.** We compared the classification accuracy performance of deep learning models with the BERT model. The accuracies of all deep learning models were high for all three datasets. Especially, for veracity datasets, all models showed >99% accuracy except BERT which showed 97.11% accuracy for veracity datasets. LSTM + CNN model showed the highest accuracy for the sentiment dataset (99.88%) while LSTM + word embeddings showed the lowest accuracy for the sentiment dataset (84.97%). For the stance dataset, LSTM + CNN showed the highest accuracy (99.96%) while BERT showed the lowest accuracy (92.57%). A detailed summary of the accuracy performance of all deep learning models for all three datasets is provided in Table 10.

#### 4.4. Limitations and Validity Threats

**4.4.1. Limitations.** One of the difficulties in building such a machine learning technique is validating many COVID-19 claims. The COVID-19 Fake News Detecting dataset, which is rather tiny, is the foundation for our approach. As a result, it may be restricted in its ability to identify new COVID-19-related disinformation. More evaluations are needed to improve the external validity of the used models. We will dig deeper into the data that we have gathered in the future and focus on higher-order metadata collecting.

**4.4.2. Validity Threats.** Every study faces a threat to validity (<https://1library.net/article/construct-validity-threats-validity-learning-structural-historical-features.z193jpvq>). Below mentioned are some of the threats to validity.

- (i) *Construct Validity* [51]: threats to construct validity centre on how theory and observation are related. This may be interpreted in our context as referring to the validity of the models that were trained and used to assess the various methodologies examined in this work. By reviewing and reevaluating with other coauthors, we have attempted to reduce any threat to construct validity. To further validate and draw conclusions from the data, more assessments are required.
- (ii) *Internal Validity* [52]: All the variables that may have influenced our results are threats to internal validity. Even though we compared the suggested process with traditional deep learning techniques

TABLE 10: Results of accuracy performance of deep learning models and BERT for sentiment, stance, and veracity datasets.

Models	Dataset	Accuracy (%)	Val accuracy (%)	Loss	Val loss
Simple RNN	Sentiment	97.80	60.49	0.0575	0.4175
	Stance	99.12	76.03	0.0790	0.2790
	Veracity	99.97	66.66	0.0309	0.3009
LSTM + WE	Sentiment	84.97	83.90	0.0871	0.0871
	Stance	96.10	89.13	0.02789	0.02789
	Veracity	99.03	94.02	0.0211	0.0211
Bidirectional + WE	Sentiment	92.5	93.08	0.0108	0.0108
	Stance	97.60	83.89	0.1709	0.1709
	Veracity	99.97	98.93	0.00928	0.00928
LSTM + CNN	Sentiment	99.88	63.08	0.0108	0.0108
	Stance	99.96	73.89	0.1709	0.1709
	Veracity	98.96	78.93	0.00928	0.00928
BERT	Sentiment	99.18	98.17	0.0308	0.0302
	Stance	92.57	96.89	0.1710	0.0959
	Veracity	97.11	99.01	0.1226	0.0210

when conducting experiments, in our context, this may apply to the training procedure described in Section 5. More comparisons, assessments, and reevaluations are required to validate the study's findings further.

- (iii) *External Validity* [51, 53]: threats to external validity concern the generalizability of the findings. Further extensive evaluations on selected and new DL models are required to reduce this threat.

## 5. Conclusion and Future Work

Since December 2019, the (COVID-19) coronavirus outbreak has aroused enormous worry within the general community and dramatically changed social attitudes and beliefs. Besides the sickness, many people suffer from anxiety and sadness from the disease. Rumours are unconfirmed facts or stories that disseminate disinformation and generate attitudes of prejudice, hatred, and fear. In this research study, we took a COVID-19 rumours dataset from news websites and tweets, combined with metadata regarding the rumours. We have collected 9200 comments from Google and 34,779 Twitter postings filtered for terms linked with COVID-19-related fake news. In experiment 1, the dataset was examined using the following three criteria: truthfulness, posture, and sentiment. In these words, we have distinct labels and performed these deep learning algorithms independently on each term. The TCN model performs best on each measurement parameter in the examined findings. So, we have adopted the TCN model for the practical implication of improved results. In experiment 2, we employed multiple state-of-the-art deep learning algorithms and examined the performance of these models on all three datasets. Based on our experimental assessments, the BERT performs better than the other state-of-the-art models assessed in our study. Some of the future directions for the fake news topic are as follows:

- (i) Rumour processing and emotional research in NLP
- (ii) Conspiracy processing and emotional study in NLP
- (iii) Recognition, prediction, rumour, and biased news categorization

- (iv) Social media network and challenging information flow study, as well as transitional network-related research

## Data Availability

The datasets presented in this study can be found in the online repository <https://github.com/MickeysClubhouse/COVID-19-rumor-dataset>.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

The authors would like to acknowledge Prince Sultan University and EIAS: Data Science and Blockchain Laboratory, for their valuable support. Also, the authors would like to thank the Prince Sultan University for funding the Article Process Charges (APC) of this publication.

## References



- [1] K. Roosa, Y. Lee, R. Luo et al., "Real-time forecasts of the COVID-19 epidemic in China from february 5th to february 24th, 2020," *Infectious Disease Modelling*, vol. 5, pp. 256–263, 2020.
- [2] L. Yan, H. T. Zhang, Y. Xiao et al., *Prediction of Criticality in Patients with Severe Covid-19 Infection Using Three Clinical Features: A Machine Learning-Based Prognostic Model with Clinical Data in Wuhan*, medRxiv, New Haven, CT, USA, 2020.
- [3] S. Bernard Stoecklin, P. Rolland, Y. Silue et al., "First cases of coronavirus disease 2019 (COVID-19) in France: surveillance, investigations and control measures, January 2020," *Euro Surveillance*, vol. 25, no. 6, Article ID 2000094, 2020.
- [4] C. Huang, Y. Wang, X. Li et al., "Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China," *The Lancet*, vol. 395, no. 10223, pp. 497–506, 2020.
- [5] E. Mahase, "Coronavirus: Covid-19 Has Killed More People than SARS and MERS Combined," *despite lower case fatality rate*, vol. 368, 2020.

- [6] T. Ai, Z. Yang, H. Hou et al., "Correlation of chest CT and RT-PCR testing for coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases," *Radiology*, vol. 296, no. 2, pp. E32–E40, 2020.
- [7] O. Gozes, M. Frid-Adar, H. Greenspan et al., "Rapid Ai Development Cycle for the Coronavirus (Covid-19) Pandemic: Initial Results for Automated Detection & Patient Monitoring Using Deep Learning Ct Image Analysis," <https://arxiv.org/abs/2003.05037>.
- [8] X. Xu, X. Jiang, C. Ma et al., "A deep learning system to screen novel coronavirus disease 2019 pneumonia," *Engineering*, vol. 6, no. 10, pp. 1122–1129, 2020.
- [9] T. Shakeel, S. Habib, W. Boulila et al., "A survey on COVID-19 impact in the healthcare domain: worldwide market implementation, applications, security and privacy issues, challenges and future prospects," *Complex & Intelligent Systems*, pp. 1–32, 2022.
- [10] H. Aldabbas, A. Bajahzar, M. Alruily, A. A. Qureshi, R. M. Amir Latif, and M. G. Farhan, "Google play content scraping and knowledge engineering using natural language processing techniques with the analysis of user reviews," *Journal of Intelligent Systems*, vol. 30, no. 1, pp. 192–208, 2020.
- [11] S. Soni and K. Roberts, "An evaluation of two commercial deep learning-based information retrieval systems for COVID-19 literature," *Journal of the American Medical Informatics Association*, vol. 28, no. 1, pp. 132–137, 2021.
- [12] A. B. Tufail, I. Ullah, R. Khan et al., "Recognition of ziziphus lotus through aerial imaging and deep transfer learning approach," *Mobile Information Systems*, vol. 2021, Article ID 4310321, 10 pages, 2021.
- [13] Y. Dai, J. Wu, Y. Fan et al., "MSEva: a musculoskeletal rehabilitation evaluation system based on emg signals," *ACM Transactions on Sensor Networks*, 2022.
- [14] R. Catelli, F. Gargiulo, V. Casola, G. De Pietro, H. Fujita, and M. Esposito, "A novel covid-19 data set and an effective deep learning approach for the de-identification of Italian medical records," *IEEE Access*, vol. 9, Article ID 19097, 2021.
- [15] J. L. Izquierdo, J. Ancochea, J. B. Soriano, S. C. R. Group, and others, "Clinical characteristics and prognostic factors for intensive care unit admission of patients with COVID-19: retrospective study using machine learning and natural language processing," *Journal of Medical Internet Research*, vol. 22, no. 10, Article ID e21801, 2020.
- [16] I. Li, Y. Li, T. Li, S. Alvarez-Napagao, D. Garcia-Gasulla, and T. Suzumura, "What are we depressed about when we talk about covid-19: mental health analysis on tweets using natural language processing," in *Proceedings of the International Conference on Innovative Techniques and Applications of Artificial Intelligence*, pp. 358–370, Springer, Cambridge, UK, December 2020.
- [17] D. Lu, D. Tong, Q. Chen, W. Zhou, J. Zhou, and S. Shen, "Exponential synchronization of stochastic neural networks with time-varying delays and Lévy noises via event-triggered control," *Neural Processing Letters*, vol. 53, no. 3, pp. 2175–2196, 2021.
- [18] A. B. Tufail, I. Ullah, W. U. Khan et al., "Diagnosis of diabetic retinopathy through retinal fundus images and 3D convolutional neural networks with limited number of samples," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 6013448, 15 pages, 2021.
- [19] A. B. Tufail, Y. K. Ma, Q. N. Zhang et al., "3D convolutional neural networks-based multiclass classification of Alzheimer's and Parkinson's diseases using PET and SPECT neuroimaging modalities," *Brain Informatics*, vol. 8, pp. 23–29, 2021.
- [20] Q. Li, L. Li, W. Wang, Q. Li, and J. Zhong, "A comprehensive exploration of semantic relation extraction via pre-trained CNNs," *Knowledge-Based Systems*, vol. 194, Article ID 105488, 2020.
- [21] J. S. Obeid, M. Davis, M. Turner et al., "An artificial intelligence approach to COVID-19 infection risk assessment in virtual visits: a case report," *Journal of the American Medical Informatics Association*, vol. 27, no. 8, pp. 1321–1325, 2020.
- [22] N. Ullah, J. A. Khan, S. Almakdi et al., "A novel CovidDetNet deep learning model for effective COVID-19 infection detection using chest radiograph images," *Applied Sciences*, vol. 12, p. 6269, 2022.
- [23] J. Roozenbeek, C. R. Schneider, S. Dryhurst et al., "Susceptibility to misinformation about COVID-19 around the world," *Royal Society Open Science*, vol. 7, no. 10, Article ID 201199, 2020.
- [24] M. Cheng, S. Wang, X. Yan et al., "A COVID-19 rumor dataset," *Frontiers in Psychology*, Article ID 644801, 2021.
- [25] H. y. Lu, C. Fan, X. Song, and W. Fang, "A novel few-shot learning based multi-modality fusion model for COVID-19 rumor detection from online social media," *PeerJ Computer Science*, vol. 7, p. e688, 2021.
- [26] T. Paul, S. Roy, S. Maity, A. Bhattacharya, S. Dutta, and S. Ghatak, "COVID-R: a deep feature learning-based COVID-19 rumors detection framework," in *Emerging Technologies in Data Mining and Information Security*, pp. 907–917, Springer, Singapore, 2021.
- [27] B. K. Yousafzai, S. A. Khan, T. Rahman et al., "Student-performulator: student academic performance using hybrid deep neural network," *Sustainability*, vol. 13, no. 17, p. 9775, 2021.
- [28] A. S. Luccioni, K. H. Pham, C. S. N. Lam, J. Aylett-Bullock, and M. Luengo-Oroz, "Ensuring the Inclusive Use of Natural Language Processing in the Global Response to COVID-19," <https://arxiv.org/abs/2108.10791>.
- [29] Q. Chen, R. Leaman, A. Allot et al., "Artificial intelligence in action: addressing the COVID-19 pandemic with natural language processing," *Annual Review of Biomedical Data Science*, vol. 4, no. 1, pp. 313–339, 2021.
- [30] T. K. Mackey, J. Li, V. Purushothaman et al., "Big data, natural language processing, and deep learning to detect and characterize illicit COVID-19 product sales: in-fovea study on twitter and instagram," *JMIR public health and surveillance*, vol. 6, no. 3, Article ID e20794, 2020.
- [31] X. Guo, H. Mirzaalian, E. Sabir, A. Jaiswal, and W. Abd-Almageed, "Cord19sts: Covid-19 Semantic Textual Similarity Dataset," <https://arxiv.org/abs/2007.02461>.
- [32] S. Verma, A. Paul, S. S. Kariyannavar, and R. Katarya, "Understanding the Applications of Natural Language Processing on COVID-19 Data," in *Proceedings of the 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, pp. 1157–1162, IEEE, Coimbatore, India, November 2020.
- [33] A. Agade and S. Balpande, *Exploring the Non-medical Impacts of Covid-19 Using Natural Language Processing 2020*.
- [34] M. Becker, S. Kasper, B. Böckmann, K. H. Jöckel, and I. Virchow, "Natural language processing of German clinical colorectal cancer notes for guideline-based treatment evaluation," *International Journal of Medical Informatics*, vol. 127, pp. 141–146, 2019.
- [35] S. A. Crossley, M. Kim, L. Allen, and D. McNamara, "Automated summarization evaluation (ASE) using natural language processing tools," in *Proceedings of the International*

- Conference on Artificial Intelligence in Education*, pp. 84–95, Springer, Chicago, IL, USA, June 2019.
- [36] T. A. Koleck, C. Dreisbach, P. E. Bourne, and S. Bakken, “Natural language processing of symptoms documented in free-text narratives of electronic health records: a systematic review,” *Journal of the American Medical Informatics Association*, vol. 26, no. 4, pp. 364–379, 2019.
- [37] R. M. Van Haren, A. M. Correa, B. Sepesi et al., “Ground glass lesions on chest imaging: evaluation of reported incidence in cancer patients using natural language processing,” *The Annals of Thoracic Surgery*, vol. 107, no. 3, pp. 936–940, 2019.
- [38] K. B. Prakash, S. S. Imambi, M. Ismail, T. P. Kumar, and Y. Pawan, “Analysis, prediction and evaluation of covid-19 datasets using machine learning algorithms,” *International Journal of Emerging Trends in Engineering Research*, vol. 8, no. 5, pp. 2199–2204, 2020.
- [39] A. Neuraz, I. Lerner, W. Digan et al., “Natural language processing for rapid response to emergent diseases: case study of calcium channel blockers and hypertension in the covid-19 pandemic,” *Journal of Medical Internet Research*, vol. 22, no. 8, Article ID e20773, 2020.
- [40] H. Abbasimehr and R. Paki, “Prediction of COVID-19 confirmed cases combining deep learning methods and Bayesian optimization,” *Chaos, Solitons & Fractals*, vol. 142, Article ID 110511, 2021.
- [41] A. B. Tufail, Y. K. Ma, M. K. A. Kaabar et al., “Deep learning in cancer diagnosis and prognosis prediction: a minireview on challenges, recent trends, and future directions,” *Computational and Mathematical Methods in Medicine*, vol. 2021, Article ID 9025470, 28 pages, 2021.
- [42] J. Wang, N. Abu-el Rub, J. Gray et al., “COVID-19 SignSym: a fast adaptation of a general clinical NLP tool to identify and normalize COVID-19 signs and symptoms to OMOP common data model,” *Journal of the American Medical Informatics Association*, vol. 28, no. 6, pp. 1275–1283, 2021.
- [43] A. B. Chapman, K. S. Peterson, A. Turano, T. L. Box, K. S. Wallace, and M. Jones, *A Natural Language Processing System for National COVID-19 Surveillance in the*, US Department of Veterans Affairs, Washington, D.C., USA, 2020.
- [44] C. Ying, Z. Huang, and C. Ying, “Accelerating the image processing by the optimization strategy for deep learning algorithm DBN,” *EURASIP Journal on Wireless Communications and Networking*, pp. 232–238, 2018.
- [45] A. S. Imran, S. M. Doudpota, Z. Kastrati, and R. Bhatra, “Cross-cultural Polarity and Emotion Detection Using Sentiment Analysis and Deep Learning—A Case Study on COVID-19,” <https://arxiv.org/pdf/2008.10031>.
- [46] K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink, and J. Schmidhuber, “LSTM: a search space odyssey,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, no. 10, pp. 2222–2232, 2017.
- [47] N. Y. Gómez-Castillo, P. E. Cajilima-Cardenaz, L. Zhinin-Vera et al., “A machine learning approach for blood glucose level prediction using a LSTM network,” in *Proceedings of the International Conference on Smart Technologies, Systems and Applications*, pp. 99–113, Springer, 2022.
- [48] C. Lea, R. Vidal, A. Reiter, and G. D. Hager, “Temporal convolutional networks: a unified approach to action segmentation,” in *Proceedings of the European Conference on Computer Vision*, pp. 47–54, Springer, Amsterdam, The Netherlands, October 2016.
- [49] R. Khan, Q. Yang, I. Ullah et al., “3D convolutional neural networks based automatic modulation classification in the presence of channel noise,” *IET Communications*, vol. 16, no. 5, pp. 497–509, 2022.
- [50] C. Lea, M. D. Flynn, R. Vidal, A. Reiter, and G. D. Hager, “Temporal convolutional networks for action segmentation and detection,” in *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, pp. 1003–1012, IEEE Computer Society, Honolulu, HI, USA, July 2017.
- [51] A. Yasin, L. Liu, T. Li, J. Wang, and D. Zowghi, “Design and preliminary evaluation of a cyber security requirements education game (SREG),” *Information and Software Technology*, vol. 95, pp. 179–200, 2018.
- [52] R. Fatima, A. Yasin, L. Liu, and J. Wang, “How persuasive is a phishing email? A phishing game for phishing awareness,” *Journal of Computer Security*, vol. 27, no. 6, pp. 581–612, 2019.
- [53] A. Yasin, L. Liu, T. Li, R. Fatima, and W. Jianmin, “Improving software security awareness using a serious game,” *IET Software*, vol. 13, no. 2, pp. 159–169, 2019.

## Research Article

# Analysis of Smart Lung Tumour Detector and Stage Classifier Using Deep Learning Techniques with Internet of Things

**Shubham Joshi** <sup>1</sup>, **Shraddha Viraj Pandit**,<sup>2</sup> **Piyush Kumar Shukla**,<sup>3</sup> **Atiah H. Almalki**,<sup>4,5</sup> **Nashwan Adnan Othman**,<sup>6</sup> **Adnan Alharbi**,<sup>7</sup> and **Musah Alhassan** <sup>8</sup>

<sup>1</sup>Department of Computer Engineering, SVKM'S NMIMS MPSTME, Shirpur Campus, Shirpur 425405, India

<sup>2</sup>Department of Artificial Intelligence & Data Science, PES Modern College of Engineering, Pune 411005, Maharashtra, India

<sup>3</sup>Department of Computer Science & Engineering, University Institute of Technology,

Rajiv Gandhi Proudyogiki Vishwavidyalaya (Technological University of Madhya Pradesh), Bhopal 462033, Madhya Pradesh, India

<sup>4</sup>Department of Pharmaceutical Chemistry, College of Pharmacy, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia

<sup>5</sup>Addiction and Neuroscience Research Unit, College of Pharmacy, Taif University, Al-Hawiyah, Taif 21944, Saudi Arabia

<sup>6</sup>Department of Computer Science, College of Science, Knowledge University, Erbil 44001, Iraq

<sup>7</sup>Department of Clinical Pharmacy, College of Pharmacy, Umm Al-Qura University, Makkah, Saudi Arabia

<sup>8</sup>University of Development Studies, Electrical Engineering Department, School of Engineering, Nyankpala Campus, Tamale, Ghana

Correspondence should be addressed to Musah Alhassan; [musahalhassan@uds.edu.gh](mailto:musahalhassan@uds.edu.gh)

Received 4 June 2022; Revised 15 July 2022; Accepted 27 July 2022; Published 13 September 2022

Academic Editor: Ateeq Ur Rehman

Copyright © 2022 Shubham Joshi et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The use of artificial intelligence (AI) and the Internet of Things (IoT), which is a developing technology in medical applications that assists physicians in making more informed decisions regarding patients' courses of treatment, has become increasingly widespread in recent years in the field of healthcare. On the other hand, the number of PET scans that are being performed is rising, and radiologists are getting significantly overworked as a result. As a direct result of this, a novel approach that goes by the name "computer-aided diagnostics" is now being investigated as a potential method for reducing the tremendous workloads. A Smart Lung Tumor Detector and Stage Classifier (SLD-SC) is presented in this study as a hybrid technique for PET scans. This detector can identify the stage of a lung tumour. Following the development of the modified LSTM for the detection of lung tumours, the proposed SLD-SC went on to develop a Multilayer Convolutional Neural Network (M-CNN) for the classification of the various stages of lung cancer. This network was then modelled and validated utilising standard benchmark images. The suggested SLD-SC is now being evaluated on lung cancer pictures taken from patients with the disease. We observed that our recommended method gave good results when compared to other tactics that are currently being used in the literature. These findings were outstanding in terms of the performance metrics accuracy, recall, and precision that were assessed. As can be shown by the much better outcomes that were achieved with each of the test images that were used, our proposed method excels its rivals in a variety of respects. In addition to this, it achieves an average accuracy of 97 percent in the categorization of lung tumours, which is much higher than the accuracy achieved by the other approaches.

## 1. Introduction

Recently, a significant number of people all across the world have become ill with the pandemic illness known as COVID-19. E-diagnosis, remote access, virtual consultants, and e-treatment have all made their way into the healthcare

industry as a result of the current climate, which has led to the elimination of the need for physical personalization and a reduction in the risk of disease transmission. It has taken the healthcare business to a new degree of severity, which has resulted in an increase in the mortality rate of people suffering from chronic illnesses, in particular those who are



afflicted with cancer, diabetes, and cardiovascular diseases. This is because there is a scarcity of medical professionals, including doctors, nurses, and radiologists [1], which has led to an increase in the number of people diagnosed with cancer. The number of people who pass away as a result of cancer and other chronic diseases is continuing to climb at an alarming rate every year throughout the whole world, especially in less developed nations. According to WHO [2], in 2010, lung cancer was the largest cause of mortality due to the disease, accounting for 1.80 million fatalities, or 18 percent of all cancer-related deaths. A disproportionate increase in transitioning countries (from 65 percent to 94 percent) versus transitioned countries (33 percent to 57 percent) is expected due to demographic change. Despite this, the global tumour burden is expected to reach 28.40 million cases in 2040, which is a 48 percent increase from 2020. This will be further exacerbated by an increase in risk factors such as smoking. Standard challenges on recognising lung tumours in patients from decade include zero symptoms that are not related to age factor, patients who have breathing problems, patients who have smoked for 30–40 years, and patients who have no symptoms [3, 4]. Numerous researchers have used a wide array of methodologies, including segmentation, detection, and classification techniques; in an effort to circumvent the challenges that are associated with the diagnosis of lung cancer [5, 6], artificial intelligence has played an increasingly since it was first introduced. This is due to the fact that artificial intelligence is suitable for solving these types of problems. The artificial intelligence-based supervised learning models that Pragma and her colleagues utilised in order to detect lung tumours and classify them as either malignant or benign can be found in [7]. The multilayer perceptron, support vector machine (SVM), and key-value network (KNN) classifiers were developed as binary lung tumour classifiers, as stated by Rodriguez et al. [8]. Dinesh and his colleagues devised a grey wolf optimization approach, which was then combined with a genetic algorithm in order to create a hybrid lung tumour classifier [9]. This was done in order to better understand how genetic algorithms work. Traditional methods, on the other hand, have a number of shortcomings, among which are the facts that they are insufficient for early detection, that they are less efficient in terms of accuracy rate, and that they are not suitable for stage categorization.

The construction of a DCNN consists of a four-layer design with a ReLU activation function [9], as can be seen above. The researchers Zhuoliu et al. created a reinforcement Q-learning system for the detection of tumours. The system classified tumours as either malignant or benign according to the stage of their growth. The author states that the challenges that arose during the process of creating the RNN model for the detection of lung cancer in terms of its localization were resolved [10].

We were encouraged by the results that these novel strategies produced, and as a result, we made the decision to use them in the healthcare industry [11]. As a result of our research, we were able to develop a new prototype algorithm that we call the Smart Lung Tumour Detector and Stage Classifier (SLD-SC). This algorithm is able to detect lung

cancer at the earliest possible stage by utilising information obtained from PET scans (Figure 1).

Our group came up with the concept for an intellectual diagnostic module that they called SLD-SC in an effort to reduce the overall mortality rate as well as to improve their ability to identify lung tumour cells caused by non-small-cell lung cancer (NSLC). Unsupervised learning methods are used in the lung cancer detector that was proposed. These algorithms are used for segmentation, feature extraction, and stage classification. In addition to the Internet of Things, there are cloud servers where databases may be stored. The diagram labelled “Figure 1” is an example of the entire system that is being discussed.

- (i) The LIDC CT DICOM images and PET scans are analysed to start, with the goal of determining the amount of noise that is now present, as well as the quantity of memory that is necessary, and so on.
- (ii) A modified version of the LSTM model has been constructed with the intention of pinpointing the parts of the lung tumour that are most specific and accurate. These zones have been segregated from one another and are now being utilised for categorization purposes in order to establish an appropriate level of lung cancer severity.
- (iii) This method makes use of multilayer convolutional neural networks, which are abbreviated as M-CNN, in order to efficiently categorise the various phases of tumour development.
- (iv) The findings of a range of tests that were carried out with the aid of numerous medical datasets are presented in this article. These experiments were carried out by the authors of this paper. In addition to this, we make use of real-time data that was obtained patients by way of an Internet of Things device.
- (v) The recommended SLD-SC has been carried out in order to carry out performance measures for a new technique, which may also be found in this work. These measurements have been carried out in this study.

The structure of the paper may be broken down into the following outline: In the second section, we looked at and spoke about the relevant literature; in the third section, we concentrated on the technique and went into further detail about it. The reasoning that underpinned this method was deconstructed and put into the appropriate context.

## 2. Related Works

Alnuaim et al. developed the unsupervised learning model known as 3D Alex Net [12]. The suggested Alex Net detection technique is put to the test using the LUNA dataset. The proposed model is inefficient because there is insufficient testing data; just 10% of the training database is being utilised. This results in an inability to accurately predict outcomes.

Tafadzwa et al. developed a supervised CNN predictor with the purpose of identifying individuals with lung cancer

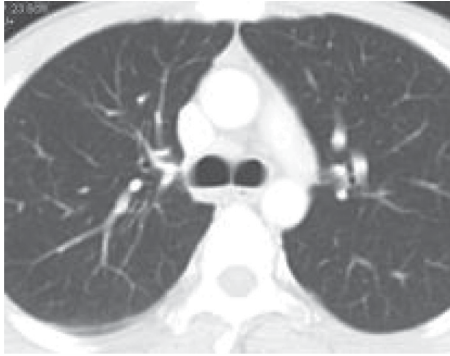


FIGURE 1: Input image.

who were in the early stages of adenocarcinoma (ADC) and squamous cell carcinoma (SCC). Validation of CNN was performed using real-time data from non-small-cell lung cancer patients obtained at Massachusetts General Hospital from patients in the early stages of the disease [13, 14].

Reference [14] According to the findings of the testing, the accuracy of the suggested approach was measured at 90.85 percent [15]. Ruoxi et al. detailed the process of determining the presence of EGFR mutations with the use of computer-assisted diagnostics. This process involves obtaining, analysing, and fusing many types of interdependent characteristics [16]. This research makes use of an innovative hybrid network model that is constructed using CNN and RNN architectural components. The CNN algorithm is used to extract the quantitative aspects of an image, and the LSTM algorithm is used to describe the connection between the various kinds of features [17].

According to their results, multitype dependency-based feature representations performed much better than single-type feature representations (accuracy of 75%, area under the curve = 0.78) when compared to traditional features that were extracted [18]. This strategy was developed in order to classify the various types of cancer that can be caused by tumour RNA sequences found in genomic data (CNN). In this particular research [19], the performance metrics that were discussed were recall, precision, and *F1*-score. According to Abdulgani et al., label-free techniques do not cause any damage to cells and do not result in any changes to the makeup of cells or their innate characteristics. The objective of this study was to enhance cell categorization by using observed optical profiles, and it did so by combining recent breakthroughs in optical measurements with Prony's techniques [20]. He and his colleagues were able to locate signature genes via the development of more accurate Tobacco Exposures Pattern (TEP) Classification models and the discovery of the interaction connections between those models on many biological levels [21]. The most current models and datasets that have been used for the execution of a variety of algorithms are summarised in Table 1.

### 3. Proposed Methodology

The proposed SLD-SC hybrid framework consists of lung tumour detector, lung tumour segmentation, and stage classifier modules. These modules allow for correct results to

be achieved in terms of “accuracy, precision, and recall.” Figure 2 provides an overview of the proposed model, which is then followed by descriptions.

#### 3.1. Outline of Proposed Work

*3.1.1. Input Image.* A PET scan was performed every two to three minutes in each of the eight or nine different bed positions. A three-dimensional iterative reconstruction approach was used to piece the photographs back together after they had been destroyed. Every individual who took part in the research was given the chance to provide their informed permission. 7 female patients and 92 male patients were present in the hospital [22]. Up to this point, there has been no study done on the differences in tumour variability between male and female NSCLC patients. As a consequence of this, while we were developing this research, we did not investigate the impact of gender on the specific features of the various cancer subtypes. It was discovered that 45 individuals were suffering from an ADC, whereas the other patients were diagnosed with SqCC. The data collection performed by the Lung Image Database Consortium (LIDC-IDRI) was responsible for providing the DICOM CT lung images [23]. Every single DICOM lung CT scan is recorded in the DICOM file format, which has a dimension of 512 by 512 pixels and is used to store the data. This data collection contains pictures of tissue slices that range from 0.45 millimetres to 0.75 millimetres in diameter and 1.15 millimetres to 2.75 millimetres in thickness. Each radiologist independently reviewed every CT scan and assigned a label to each lesion, based on which of the following three groups it belonged to: nodules, seminodules, or nonnodules. The different forms of tumours were determined after the examination of the CT images by four radiologists [24]. Table 2 represents the patient ID with stages.

Selected nodules were characterised as “well-circumscribed, juxta-vascular, juxta-pleural, pleural-tail.” The ACM method centred its attention on a location of interest (a lung tumour) that had a nodule size of (75 × 85 × 45) millimetres [25]. It is possible to calculate a few radiologists stated the number of voxels that were engaged across all dimensions and extract information on the size of the nodule, as shown in Table 2. The features of both the genuine and the simulated nodules are mentioned in Table 2.

#### 3.2. Lung Tumor Detection Using Modified LSTM Model.

Applications that need text recognition and voice processing are probably the ones that make use of RNN because of the storage capacity it offers [11]. The process of transferring data is carried out from one state to the next inside the network in a sequential method. The limitations of this recurrent neural network include making it inappropriate for lengthy sequence prediction, and it also suffers from vanishing gradient corruption [26]. Connected cells and gates form the LSTM's memory block, which is made up of the LSTM's building blocks. These gateways and cells are used rather often for the goals of retaining input states and updating [27]. Table 3 represents the LSTM features [28].

TABLE 1: Existing methodology with comparison of the performance metrics.

Paper details	Techniques used in the existing methodology	Datasets available	Accuracy rate of the existing work
Tafadzwa et al. (2021)	Supervised CNN predictor	LUAD	AUC = 71%
Pragya et al. (2021)	SVM, KNN, and CNN	LIDC-IDRI, LUNA 16	Accuracy = 91%
Kalaivani et al. (2021)	Deep CNN model	LIDC-IDRI	Accuracy = 90.85%
Khalifa et al.	BPSO-DT	LUNA	Acc = 88.25%
Abdulgani et al.	TEP classification model	LUAD	Accuracy = 92.65

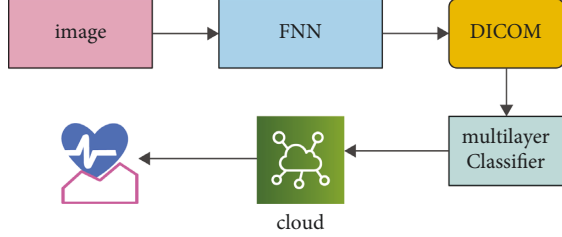


FIGURE 2: Proposed work framework.

TABLE 2: Patient ID with stages.

Patient ID	Stage
LUNG1-001	2
LUNG1-002	2
LUNG1-003	2
LUNG1-004	2
LUNG1-005	4
LUNG1-006	3

LSTM performs better than conventional RNNs due to the distinctive nature of its memory. The structure of an LSTM is shown below, along with a mathematical representation of it. The memory blocks are separated into three categories [29]: (i) the initial loading gate (), (ii) the final output gate (), and (iii) the middle inspection gate (), with equation 3 representing the activation function.

$$\alpha(a) = \frac{1}{1 + e^{-a}}, \alpha(a) \in \{0, 1\}. \quad (1)$$

The architecture of LSTM blocks may be observed in Figure 3, along with the training function, which is indicated by  $q_N$ , and used for executing the equation that describes the entrant states and block values. The kernel functions, which are defined by (6) and (7), are as follows. The structure of the model is represented as a combination of three blocks, and this combination is indicated as “E N (E blocks; )”. The entrant signals of the initial blocks are followed by various states once the model has been constructed. Between the two different units,  $u$  and  $r$ , the weights of the strata are denoted by the notation “ $Wu, r$ ”. When the loading signals are interfaced with the gates at time ‘ $j$ ’, the resulting signal, which also includes the outcome and the feedback signals, is indicated as  $EN(j)$ . The indications of the various entrance places are listed down below.

$$E_N(j) = (q_N(j)), N \in \{\delta, \alpha, \emptyset\}, \quad (2)$$

$$P_N(j) = \sum_y^x zW_N * E_N(j), N \in \{\delta, \alpha, \emptyset\}. \quad (3)$$

TABLE 3: LSTM features.

Layer 1	Layer 2	Layer $n$
64	64	64
256	256	256

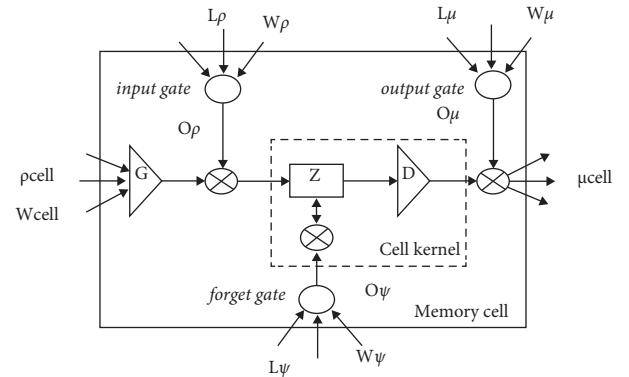


FIGURE 3: Structure of LSTM network.

The entrant block is represented as “ $\delta$  block ( $j-1$ )” and as “ $W$ block ( $j-1$ )” called weights and equation.

$$q_{C_v^h}(j) = \sum_y^x zWblock * \delta blocks(j). \quad (4)$$

The product of incoming signals from cells with loading and overlook gates, as well as previous state information, is used to calculate the internal state of the blocks.

$$RC_v^h(j) = \left\{ \begin{array}{l} 0 \quad t = 0, \emptyset_{a,k}(j) * RC_v^h(j-1) \\ + \emptyset_{in,k}(j) \cdot E(q_{C_v^h}(j)), j > 0 \end{array} \right\}, \quad (5)$$

$$\gamma(a) = \frac{4}{1 + e^{-a}} - 2 \\ = 2 \cdot \tanh \tanh\left(\frac{a}{2}\right), \gamma(a) \in [-2, 2]. \quad (6)$$

“ $\vartheta_{C_v^h}(j)$ ” denotes the output cell block “ $C_v^h$ ” at the sampling time “ $j$ ” which is calculated as follows:

$$\vartheta_{C_v^h}(j) = \emptyset_{out,j}(j) * (\gamma_{C_v^h}(j)), \quad (7)$$

$$\alpha(a) = \tanh\left(\frac{a}{2}\right), h(a) \in [-1, 1]. \quad (8)$$

DicomCT lung imaging makes it more difficult to detect nodules in the lungs than it does in other organs, such as the brain. This makes it more challenging to identify lung nodules.

- (1) If the lung parenchyma was not successfully restored, the patient's mediastinum and thoracic wall will need to be removed from their thorax.

Step two involves the use of the Active Contour Model to segment the part of the lung picture that represents the tumour (ACM) [30].

**3.3. Lung Tumor Segmentation.** The data were used to generate a curve that aids in the identification of tumour sections in the relevant photographs [31]. This curve was then applied to the data. The Snake model was used to create the curve. It is necessary to begin by drawing the curve around the item that has been provided, and then it is necessary for the curve to move its location towards the interior of the object before coming to an end at the object's limits [32]. The proposed method generates three-dimensional features for the CNN classifier by combining two-dimensional stochastic characteristics with three-dimensional data [33]. These three-dimensional features are then fed into the CNN classifier. The segmentation image that was produced from a PET scan as a result of the segmentation method is shown in Figure 4. In the dataset provided by LIDC-IDRI, we were successful in separating out the tumour component, as can be shown in Figure 5.

**3.4. Input Validation of the Proposed Work.** As can be seen in Figure 6, the authors of this study make heavy use of the 10-fold cross-validation technique for the LIDC-IDRI/PET dataset. This study made use of the dataset after it had been randomly segmented into stratified 10-fold cross validation. The dataset is one that has been utilised in a great deal of research that is based on deep convolutional neural networks as well as traditional machine learning techniques applied to bioimages and biosignals [34]. A high-performance computing system was developed by the use of a personal computer equipped with an NVIDIA GeForce GTX 1650, a Deep CNN model that was trained from the ground up, along with some random weight (HPC). The number of epochs that will be used for each training phase will be 150 from the dataset that has been provided, and early stopping procedures have been used so that the dataset does not get overfit [35]. In addition, the accuracy analysis with relation to epochs is shown in a very straightforward manner in Figure 6, which covers both the training data and the validation data.

Figure 7 shows the performance metrics of the proposed work.

**3.5. Lung Tumor Stage Classification.** The human brain is used to stimulate artificial neural networks, which enables machine learning to be applied to the solution of complicated issues [36]. Deep learning is one of the subfields that

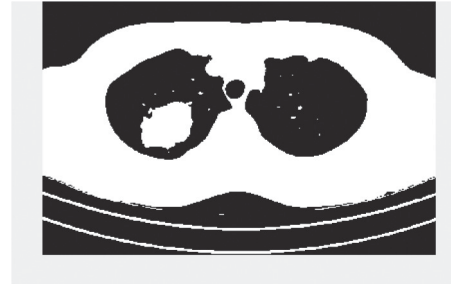


FIGURE 4: CT image after preprocessing.

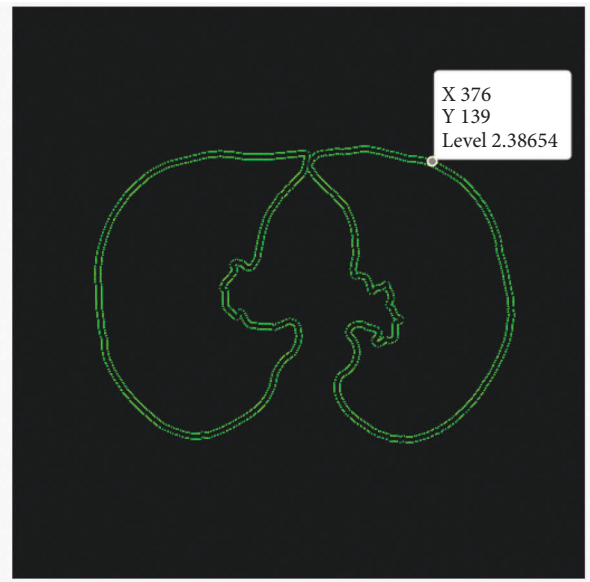


FIGURE 5: Segmented input image dataset.

fall under the umbrella of machine learning (DL). Figure 8 shows the classification-based image database.

The DL method is used for the task of extracting characteristics from massive volumes of data; using DL algorithms to glean useful information from massive amounts of data is advantageous in a number of different ways [36]. Because identifying a feature takes a significant amount of time and may be quite expensive, learning-related applications of DL methods do not need the usage of labelled data in any way. In the context of healthcare, we could have both labelled and unlabelled kinds of data, such as X-ray photos taken regardless of the patient's medical state [37], enormous volumes of data that are not labelled. Figure 9 shows the basic CNN structure.

There are various different deep learning strategies available for your selection. In this part of the article, we discussed some of the most often used examples among them. (1) An artificial neural network, also known as an ANN, is a method for deep learning that consists of a multitude of hierarchical layers and uses perceptrons [38], which are essentially neurons, as its fundamental building pieces. There are many layers used, starting with the input layer and going to the hidden layer, which performs the functions of both the training layer and the output layer. The

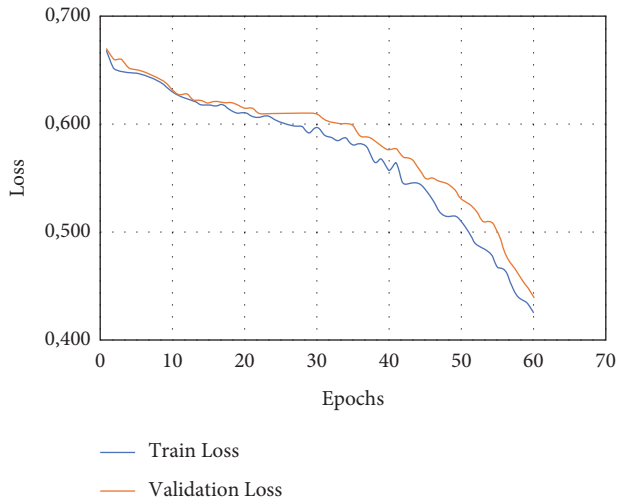


FIGURE 6: Overall proposed structure performance analysis 1.

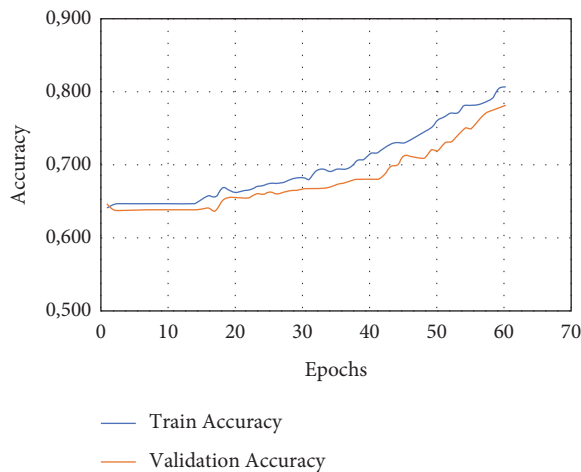


FIGURE 7: Overall analysis of performance.

process begins with the input layer. It is conceivable that the results will not be improved even if the number of concealed layers is greatly increased. This is one of the possibilities [39]. Overfitting may also occur if a high number of layers are added all at once. This can lead to a huge number of distortions, which can provide an excessive amount of interference in the data that is being collected. This can cause overfitting. Up until this moment, the convolutional neural network, also known as a CNN (Figure 7(a)), has been the most recognised example of a useful technology in the field of healthcare. A flight path that has a static extent is being used here as an input [40]. It is possible to utilise it to process medical data, such as image processing for the diagnosis of lung tumours, for instance. It is feasible to connect a large number of perceptrons together and give each of them a weight that is capable of being changed after each iteration of the algorithm [41]. A network is said to be feedforward if waves only go in one direction through it, from the input layer all the way to the output layer. This means that the feedforward network only has one direction in which waves may travel [42].

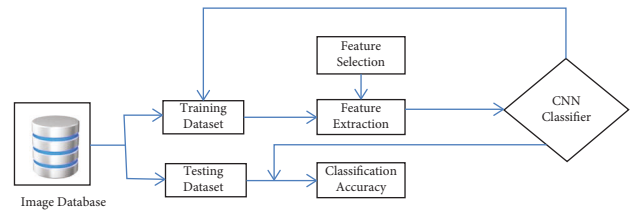


FIGURE 8: Classification based on image database.

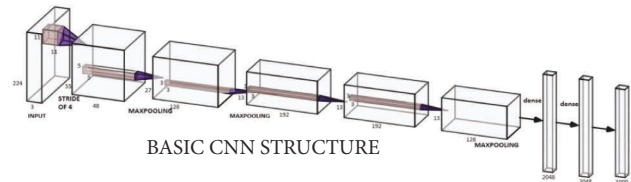


FIGURE 9: Basic CNN structure.

Figure 10 shows the architecture of proposed technique. It is also one of the deep learning approaches that is used the most frequently. The fact that it is classified as a feedforward network while having several layers indicates that its operation proceeds in just one direction, namely, from input to output [43]. The transmission of data via the process layer results in the extraction of useful features from the input data, which are then shown in the output layers as a direct consequence of the extraction. In the field of medicine, it is used to the process of diagnosing sickness based on samples of tissue collected from patients. Read structures, which are sometimes tough to interpret by human medical professionals, are encouraged to be used as a result of this. Figure 10 is a representation of the proposed architecture for the CNN, which, when implemented, would result in the classification of lung tumour stages based on the data obtained [44]. Table 4 shows the tumor stages.

When compared to other types of tumours, the staging method that is used for lung tumours is what distinguishes them [45]. The techniques that are used for determining the stage of a lung tumour are, for the most part, defined by the specific experiences of doctors as well as the general agreement of the medical community found in both Tables 4 and 5.

**3.6. IoMT Framework.** An IoMT-based lung tumour identification and stage classification approach is discussed in this part of the study. This methodology makes use of deep learning algorithms to predict the lung tumour picture that has been provided by the public and an end user. Figure 11 shows the Internet of things architecture.

The Internet of Things (IoT) devices are networked with one another in the field of computer vision for the purpose of data transfer across a network. This calls for interaction between people and computers, in addition to contact between individuals on the human side. The connection provided by the Internet of Things (IoT) is seeing meteoric expansion in the realm of healthcare. This endeavour will not only be useful to patients, but also to the doctors who treat

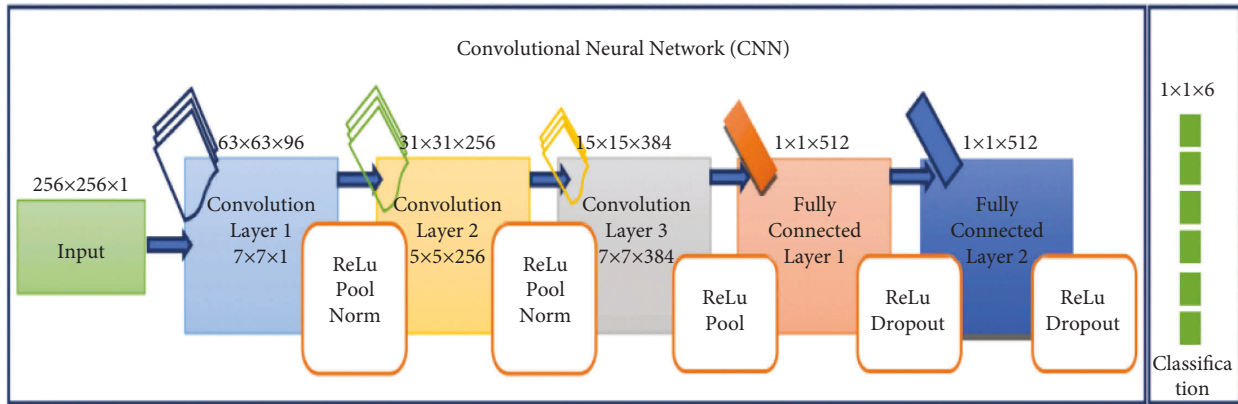


FIGURE 10: Architecture of the proposed technique.

TABLE 4: Lung tumor stages classified based on its size.

Features	Extracted
Standard deviation	0.12346
Mean	0.24597
Median	0.36798
Entropy	0.46479
Skewness	0.89764

TABLE 5: Estimated classification results for test data.

Images	Classification trained	Classification tested
Img 1	Class 1	Class 1
Img 2	Class 2	Class 2
Img 3	Class 1	Class 1
Img 4	Class 1	Class 1
Img 5	Class 2	Class 2
Img 6	Class 1	Class 1
Img 7	Class 2	Class 1

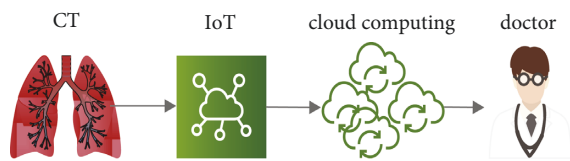


FIGURE 11: Virtual monitoring and E-diagnosis framework.

those patients. For instance, with the assistance of a connection to the Internet of Things, patients can get preventative advice from their doctor without having to see him or her, and they can also submit real-time data to their doctor for improved treatment without having to pay a visit. Both of these benefits are made possible without the patients having to physically see their physician. Both of these advantages are available to patients without the prerequisite of their having to schedule an appointment. Figure 11 illustrates how data may be gathered and transferred between devices that are located in different locations by using sources that are accessible, such as networks and sensors. The diagram is a useful tool for illustrating this procedure. The Internet of Things (IoT) is an essential technology that at the moment

has the potential to be employed in the administration of remote medical care. The Internet of Things (IoT) is a network of linked devices that may be worn or implanted and are powered by lightweight and tiny batteries. These devices may also communicate with one another through the Internet. It distributes the information that was gathered by sensors and transmitted through the network to medical institutions, such as hospitals and clinics, in addition to sending it to other sites. It is vital that the data be safeguarded while simultaneously making it accessible to all relevant parties since the existence of this data, which is developing at a rapid pace and may be referred to as “big data,” making it necessary for the data to be protected. The data will be shared in an efficient way by making use of an intelligent and secure architecture, which will be put into place at a variety of medical institutions. This will ensure that the data is kept confidential. Figure 4 depicts the comprehensive structure of the system as a whole, which is based on block chains and makes use of cloud storage to store electronic medical records (EMRs) and other types of information. The information that was acquired by instruments is first delivered to a PDA device under the architecture that we propose. After that, the PDA device will produce the hash of the medical information by making use of the typical hash algorithms. At this point, the hash will be uploaded to a private block chain link that is accessible through the Internet. Everyone involved in the healthcare industry, like as hospitals, insurance companies, laboratories, and other organisations, will function as a block-chain node to enable the free flow of information. This includes patients, doctors, and lab technicians. This includes hospitals, clinics, and research centres as well as other types of labs. Hash is being directed by a PDA device; it will be recognised by every node in the network. This data block has to be validated and verified with the help of the nodes so that the hash can be identified. The authentication operation is carried out on the basis of the hash that has already been established, and after that, as is customary, the hash of the most recent data block is compared to the hash that has already been established. It is possible that the hash of the most recent block of data generated by the PDA device will be included in the block of data that is produced by the PDA device. This is something that is a possibility. There is a possibility that this will take

place. The overwhelming majority of the nodes that comprised the block chain required authenticating blocks before they could participate.

Following the completion of the approval process, the block will be placed to the queue, at which time it will be assigned an identification in the form of an ID number and a concealed key that cannot be replicated will be generated. After having been sent earlier, the ID and key are then received by the PDA device where they were previously delivered. Using the key, the PDA device encrypts the real medical data and then sends the encoded data, together with the ID and hash of the health data, to the cloud-based database server so that it can be further processed. The ID and hash of the health-data are used in the verification process to ensure that the data are genuine. If someone makes an attempt to corrupt the data stored in a single block, then that attempt will have an impact on the data stored in the blocks that follow it in the chain. After the data has been recognised by using the ID and decoding has been completed by using the key, certain medical institutions may make a request to access medical information that has been compiled and stored in the cloud storage. This request could come after the data has been recognised using the ID. Patients will gain access to the previously encrypted medical information as soon as the decoding procedure is finished. As a result of the fact that a great number of apps for medical care that are powered by deep learning employ this answer as the default, it is essential to be familiar with it. In the realm of healthcare, deep learning systems have a great many different applications that they may do. Several of these applications, such as Medical Assessment Provision, Modified Managements, and Predictive Healthcare, amongst others, have been discussed in the past. The fact that this architecture facilitates the safe transfer of data across a large number of health institutions and that the data may be used in a diverse collection of applications that are pertinent to the field of medical care is the primary advantage that this architecture provides. In our design, the prerequisites for adequate safety have been satisfied. Authenticated users are the only ones who can see and save data, whereas all other users are unable to do so. Figure 12 shows the IoMT framework.

Techniques using cryptography are helpful in the safeguarding of confidential information. Both the design of the block-working chain and the fact that the access-regulator is centred on a secret key contribute to the fact that the system's integrity is not compromised in any way. Due to the fact that the hidden key can only be created by the nodes of the block chain, information can only be decrypted by those nodes. It is also important to note that basic safety measures contributed to the value of IoMT and cloud-supported medical care structure advantages in order to prevent other calculating difficulties that result in extra resource consumption as a result of the independent execution of encryption algorithms. This is important to note because it is important to note that basic safety measures contributed to the value of IoMT and cloud-supported medical care structure advantages. If the strategy of encrypting information at the source and then decrypting it at the matching

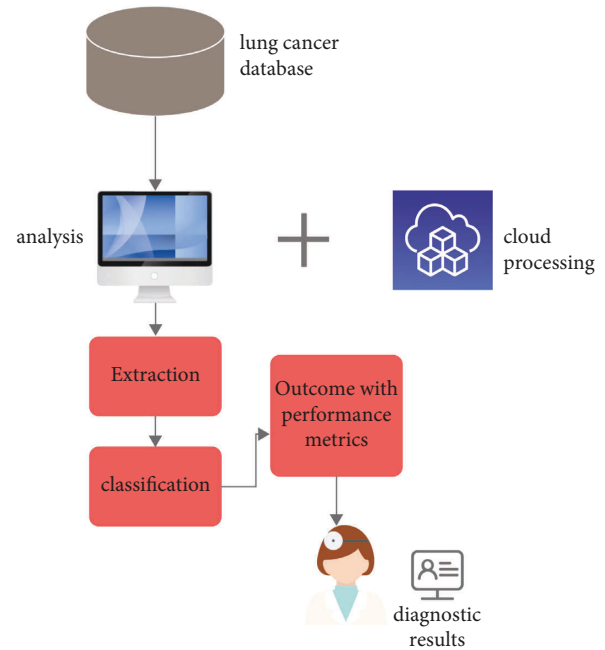


FIGURE 12: IoMT framework.

destination is employed, it will take a lengthy time, which is unacceptable in applications such as health care. This is because medical information is accessible by a large number of parties.

#### 4. Analysis with the Experimental Outcomes

Utilizing several learning modules, the framework is constructed inside of an integrated development environment (IDE) written in *Python*. For the objectives of pretraining and evaluation, both LIDC-IDRI scanned images and PET scanned photos are used in conjunction with one another. The platforms that have been used to verify these models are as follows: “NVIDIA JETSON GPU System-on-Module with 256-core NVIDIA Pascal™ GPU architecture and 256 NVIDIA CUDA cores” and “NVIDIA JETSON GPU System-on-Module with 256-core NVIDIA Pascal™ GPU architecture and 256 NVIDIA CUDA cores.” When compared to other models that are presently being used, the model that has been presented has a lower level of temporal complexity. This can be attributed to the model's optimal speed as well as its efficient design. This is due to the fact that the model that was recommended was created. Applying the standard formula, which is presented for your reading in the following part for your convenience, is what is done to determine the performance parameters.

Figure 13 depicts a comparative comparison of our proposed model with current models, such as the SVM model and the enhanced FCM model, using performance indicators, such as accuracy, sensitivity, and specificity.

In Figure 14, TS-1 to TS-5 describe the types of stages in lung cancer that is predicted correctly and misclassified. It shows that the TS-1 can be found at early stages with 88.56% accuracy rate.

		Predicted Class		
		Positive	Negative	
Actual Class	Positive	True Positive (TP)	False Negative (FN) Type II Error	Sensitivity $\frac{TP}{(TP + FN)}$
	Negative	False Positive (FP) Type I Error	True Negative (TN)	Specificity $\frac{TN}{(TN + FP)}$
		Precision $\frac{TP}{(TP + FP)}$	Negative Predictive Value $\frac{TN}{(TN + FN)}$	Accuracy $\frac{TP + TN}{(TP + TN + FP + FN)}$

FIGURE 13: Estimated performance metrics for validating the proposed SLD-SC model.

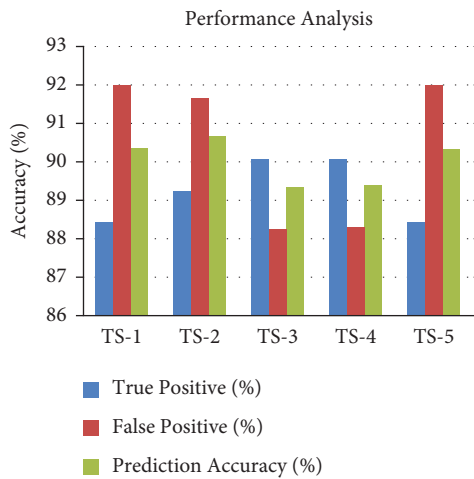


FIGURE 14: Performance analysis of TP, FP, and accuracy.

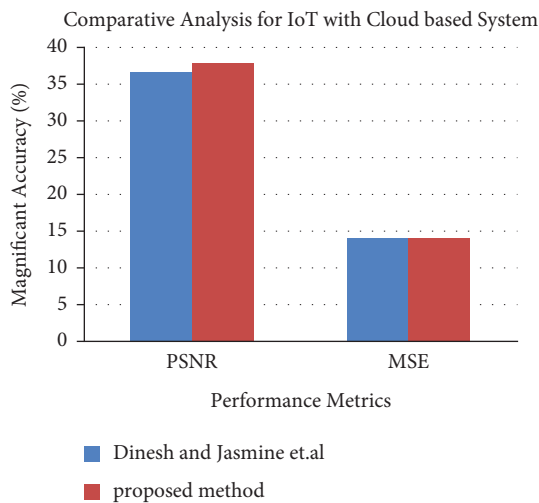


FIGURE 15: Performance metrics for IoT with Cloud-based system.

Figure 15 shows the results of a comparison between the proposed technique with the performance metrics of existing IoMT techniques are shown in Figure 15. This is as a result of the utilisation of superior lung tumour segmentation techniques for more accurate tumour prediction selection.

### 5. Conclusion

This article focuses on the construction of a lung tumour detector that is based on the Internet of Things in order to decrease the mortality rate that is linked with lung tumours. The suggested SLD-SC module was able to identify the tumour with a higher degree of precision thanks to the use of unsupervised learning neural networks both as a predictor and stage classifier. The LSTM model, which is the most popular Q-learning model, has had its structure updated, and it has been constructed as a lung tumour detector based, among other things, on characteristics retrieved from DICOM and PET scanned pictures. This was accomplished by using the information obtained from those two types of images. On the basis of the segmented photos, which are subsequently fed back into the segmenter, a multilayer CNN model is used to identify and grade each patient's severity level. This information is then sent back into the segmenter. For the segmented images that were put through the evaluation process, an accuracy level of 97 percent in stage categorization was achieved. These deep learning and smart lung tumour models are able to identify tumours at an earlier stage than was previously possible, and this is made possible by the use of superior virtual monitoring and E-diagnosis tools. During the course of our future work, we want to first construct and then analyse the suggested system for huge databases. Following that, we will publish our results about how the system is able to manage vast amounts of hospital data in a secure manner. In addition to that, it is intended that throughout the course of future development, this prototype will be turned into a fully functional Internet of



Things hardware. Limitations of the proposed work include only few characteristics that have been retrieved for cancer nodules. No preprocessing like noise reduction and picture smoothing which might possibly help in boosting the detection of nodules correctly has been applied. No categorization as benign or malignant of removed cancer has been conducted [46–49].

## Data Availability

The data that support the findings of this study are available at <https://www.kaggle.com/datasets/mohamedhanyyy/chest-ctscan-images>.

## Conflicts of Interest

The authors declare that they have no conflicts of interest related to this work.

## Acknowledgments

This study was supported by University Researchers Supporting Project no. TURSP-2020/208, Taif University, Taif, Saudi Arabia.




## References

- [1] S. Singh Yadav, M. Zain, P. Sahai, and S. Porwal, “Vandana chauhan “challenges encountered in cancer care and management during covid-19 in south asian countries”,” *Asian pacific journal of cancer care*, vol. 5, 2020.
- [2] World Health Organization, *Coronavirus Disease 2019 (COVID-19) Situation Report*, Technical Report March, World Health Organization, Geneva, Switzerland, 2020.
- [3] A. A. Alnuaim, M. Zakariah, P. K. Shukla et al., “Human-computer interaction for recognizing speech emotions using multilayer perceptron classifier,” *Journal of Healthcare Engineering*, vol. 202212 pages, Article ID 6005446, 2022.
- [4] H. Bhardwaj, P. Tomar, A. Sakalle, D. Acharya, T. Badal, and A. Bhardwaj, “A DeepLSTM model for personality traits classification using EEG signals,” *IETE Journal of Research*, pp. 1–9, 2022.
- [5] D. E. Jonas, D. S. Reuland, S. M. Reddy et al., “Screening for lung cancer with low-dose computed tomography,” *JAMA*, vol. 325, no. 10, pp. 971–987, 2021.
- [6] Z. Xiao, B. Liu, L. Geng, F. Zhang, and Y. Liu, “Segmentation of lung nodules using improved 3D-UNet neural network,” *Symmetry*, vol. 12, no. 11, p. 1787, 2020.
- [7] A. A. Alnuaim, M. Zakariah, A. Alhadlaq et al., “Human-computer interaction with detection of speaker emotions using convolution neural networks,” *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 7463091, 16 pages, 2022.
- [8] A. Sakalle, P. Tomar, H. Bhardwaj et al., “Genetic programming-based feature selection for emotion classification using EEG signal,” *Journal of Healthcare Engineering*, vol. 2022, 6 pages, 2022.
- [9] M. Savic, Y. Ma, G. Ramponi, W. Du, and Y. Peng, “Lung nodule segmentation with a region-based fast marching method,” *Sensors*, vol. 21, no. 5, p. 1908, 2021.
- [10] K. W. Doo, E. Y. Kang, H. S. Yong, O. H. Woo, K. Y. Lee, and Y. W. Oh, “Accuracy of lung nodule volumetry in low-dose CT with iterative reconstruction: an anthropomorphic thoracic phantom study,” *British Journal of Radiology*, vol. 87, no. 1041, Article ID 20130644, 2014.
- [11] P. Chaturvedi and A. Jham, “Meet vanani and VarshaNemad “prediction and classification of lung cancer using machine learning techniques”,” *IOP Conference Series: Materials Science and Engineering*, vol. 1099, 2021.
- [12] A. A. Alnuaim, M. Zakariah, C. Shashidhar et al., “Speaker gender recognition based on deep neural networks and ResNet50,” *Wireless Communications and Mobile Computing*, vol. 202213 pages, Article ID 4444388, 2022.
- [13] D. Valluru and I. Jasmine SelvakumariJeya, “IoT with cloud based lung cancer diagnosis model using optimal support vector machine,” *Health Care Management Science*, vol. 23, 2019.
- [14] M. Fatan Serj, B. Lavi, G. Hoff, and D. P. Valls, “A Deep Convolutional Neural Network for Lung Cancer Diagnostic,” *Computer Vision and Pattern Recognition*, 2018, <https://arxiv.org/abs/1804.08170>.
- [15] Z. Liu, G. Zhang, Z. Jingyuan et al., “Second-generation sequencing with deep reinforcement learning for lung infection detection,” *Journal of Healthcare Engineering*, vol. 2020, Article ID 3264801, 9 pages, 2020.
- [16] M. L. R. AbdElNabi, M. Wajeeh Jasim, H. M El, M. Hamed N Taha, and N. E. M Khalifa, “Breast and colon cancer classification from gene expression profiles using data mining techniques,” *Symmetry*, vol. 12, no. 3, p. 408, 2020.
- [17] R. Zhang, C. Cheng, X. Zhao, and X. Li, “Multiscale mask R-CNN-based lung tumor detection using PET imaging,” *Molecular Imaging*, vol. 18, Article ID 153601211986353, 2019.
- [18] E. S. Neal Joshua, D. Bhattacharyya, M. Chakkravarthy, and Y. C. Byun, “3D CNN with visual insights for early detection of lung cancer using gradient-weighted class Activation,” *Journal of Healthcare Engineering*, vol. 2021, Article ID 6695518, 11 pages, 2021.
- [19] A. Purohit, A. Bhardwaj, A. Tiwari, and N. S. Choudhari, “Removing code bloating in crossover operation in genetic programming,” in *Proceedings of the 2011 International Conference on Recent Trends in Information Technology (ICRTIT)*, pp. 1126–1130, IEEE, Chennai, India, June 2011.
- [20] P. K. Shukla, M. Zakariah, W. A. Hatamleh, H. Tarazi, and B. Tiwari, “AI-DRIVEN novel approach for liver cancer screening and prediction using cascaded fully convolutional neural network,” *Journal of Healthcare Engineering*, vol. 2022, Article ID 4277436, 14 pages, 2022.
- [21] T. L. Chaunzw, H. Ahmed, A. S. YiwenXu et al., “Deep Learning Classification of Lung Cancer Histology Using CT Images,” *Scientific Reports*, vol. 11, 2021.
- [22] Y. Pathak, P. K. Shukla, and K. V. Arya, “Deep bidirectional classification model for COVID-19 disease infected patients,” *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 18, no. 4, pp. 1234–1241, 1 July-Aug. 2021.
- [23] A. F. Al Ahmad and M. Al Ahmad, “Label-free normal and cancer cells classification combining prony’s method and optical techniques,” *IEEE Access*, vol. 8, Article ID 32882, 2020.
- [24] A. U. Rehman, A. Jiang, A. Rehman, A. Paul, S. Din, and M. Tariq Sadiq, “Identification and Role of Opinion Leaders in Information Diffusion for Online Discussion Network,” *Journal of Ambient Intelligence and Humanized Computing*, 2020.
- [25] Q. He, Z. Qiu, Y. Song, and K. Song, “A new TTZ feature extracting algorithm to decipher Tobacco related mutation signature genes for the personalized lung adenocarcinoma treatment,” *IEEE Access*, vol. 8, Article ID 89031, 2020.

- [26] A. U. Rehman, R. Tariq, A. Rehman, and A. Paul, "collapse of online social networks: structural evaluation, open challenges, and proposed solutions," in *Proceedings of the 2020 IEEE globecom workshops (GC wkshps)*, pp. 1–6, Taipei, Taiwan, December 2020.
- [27] N. Kalaivani, N. Manimaran, S. Sophia, and D. D. Devi, "Deep learning based lung cancer detection and classification," *IOP Conference Series: Materials Science and Engineering*, vol. 994, 2020.
- [28] G. Sinha, *Deep Learning Method for Object Detection: R-CNN Explained* towards Data Science, Canada, 2020.
- [29] T. A. Ahanger, A. Alqahtani, M. Alharbi, and A. Algashami, "Cognitive decision-making in smart police industry," *The Journal of Supercomputing*, vol. 78, no. 10, Article ID 12834, 2022.
- [30] B. Khan, P. K. Shukla, M. K. Ahirwar, and M. Mishra, "Strategic analysis in prediction of liver disease using different classification algorithms," in *Handbook of Research on Disease Prediction Through Data Analytics and Machine Learning*, G. Rani and P. Tiwari, Eds., IGI Global, Hershey, Pennsylvania, pp. 437–449, 2021.
- [31] R. Ali Naqvi, D. Hussain, and W. K. Loh, "Artificial intelligence-based semantic segmentation of ocular regions for biometrics and healthcare applications," *Computers, Materials & Continua*, vol. 66, no. 1, pp. 715–732, 2020.
- [32] R. Xi Qin, Z. Wang, L. Jiang et al., "Fine-grained lung cancer classification from PET and CT images based on multidimensional attention mechanism," *Complexity*, vol. 2020, Article ID 6153657, 12 pages, 2020.
- [33] A. Masood, B. Sheng, P. Li et al., "Computer-assisted decision support system in pulmonary cancer detection and stage classification on CT images," *Journal of Biomedical Informatics*, vol. 79, pp. 117–128, 2018.
- [34] H. Sak, A. Senior, and F. Beaufays, *Long Short-Term Memory Recurrent Neural Network Architectures for Large Scale Acoustic Modeling*, Google, USA, 2014.
- [35] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, 1997.
- [36] S. J. GopiKasinathan and A. H. Gandomi, "Automated 3-D lung tumor detection and classification by an active contour model and CNN classifier," *Expert Systems with Applications*, vol. 134, 2019.
- [37] R. Rouhi, M. Jafari, S. Kasaei, and P. Keshavarzian, "Benign and malignant breast tumors classification based on region growing and CNN segmentation," *Expert Systems with Applications*, vol. 42, no. 3, pp. 990–1002, 2015.
- [38] H. Kaushik, D. Singh, M. Kaur, H. Alshazly, A. Hamam, and H. Hamam, "Diabetic retinopathy diagnosis from fundus images using stacked generalization of deep models," *IEEE Access*, vol. 9, Article ID 108276, 2021.
- [39] K. Gopi and J. Selvakumar, "Analysis of lung tumour detection and segmentation using level set method of active contour model," *International Journal of Engineering & Technology*, vol. 7, pp. 410–412, 2018.
- [40] Y. J. Peng and S. Zhang, "A novel approach for active contour initialization in 3D medical image segmentation," in *Applied Mechanics and Materials*, vol. 239, pp. 1004–1010, Trans Tech Publ, 2013.
- [41] D. Poap, M. Wozniak, R. Damaševičius, and W. Wei, "Chest radiographs segmentation by the use of nature-inspired algorithm for lung disease detection," in *Proceedings of the 2018 IEEE Symposium Series on Computational Intelligence (SSCI)*, pp. 2298–2303, Bangalore, India, November 2018.
- [42] M. A. Khan, V. Rajinikanth, S. C. Satapathy et al., "VGG19 network assisted joint segmentation and classification of lung nodules in CT images," *Diagnostics*, vol. 11, no. 12, p. 2208, 2021.
- [43] D. Singh, M. Kaur, M. Y. Jabarulla, V. Lee, and H.-N. Lee, "Evolving fusion-based visibility restoration model for hazy remote sensing images using dynamic differential evolution," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–14, 2022.
- [44] T. A. Ahanger, "IoT Inspired Smart Environment for Personal Healthcare in Gym," *Neural Computing and Applications*, vol. 62, 2022.
- [45] N. U. Rehman, M. S. Zia, T. Meraj et al., "A self-activated cnn approach for multi-class chest-related covid-19 detection," *Applied Sciences*, vol. 11, no. 19, p. 9023, 2021.
- [46] M. Bicakci, O. Ayyildiz, Z. Aydin, A. Basturk, S. Karacavus, and B. Yilmaz, "Metabolic imaging based sub-classification of lung cancer," *IEEE Access*, vol. 8, Article ID 218470, 2020.
- [47] M. M. Wille, A. Dirksen, H. Ashraf et al., "Results of the randomized Danish lung cancer screening trial with focus on high-risk profiling," *American Journal of Respiratory and Critical Care Medicine*, vol. 193, no. 5, pp. 542–551, 2016.
- [48] A. Sakalle, P. Tomar, H. Bhardwaj, D. Acharya, and A. Bhardwaj, "An analysis of machine learning algorithm for the classification of emotion recognition," in *Soft Computing for Problem Solving*, pp. 399–408, Springer, Singapore, 2021.
- [49] D. Sharma and G. Jindal, "Computer aided diagnosis system for detection of lung cancer in CT scan images," *International Journal of Computer and Electrical Engineering*, vol. 3, no. 5, pp. 714–718, 2011.

## Research Article

# Performance Evaluation of Bio Concrete by Cluster and Regression Analysis for Environment Protection

Ashish Shukla,<sup>1</sup> Nakul Gupta,<sup>1</sup> Kunwar Raghvendra Singh,<sup>1</sup> Pawan Kumar Verma ,<sup>2</sup> Mohit Bajaj ,<sup>3,4</sup> Arfat Ahmad Khan ,<sup>5</sup> and Frie Ayalew <sup>6</sup>

<sup>1</sup>Department of Civil Engineering, GLA University, Mathura, Uttar Pradesh 281406, India

<sup>2</sup>Department of Computer Science Engineering, MIT Art, Design and Technology University, Pune, Maharashtra 412201, India

<sup>3</sup>Department of Electrical Engineering, Graphic Era (Deemed to Be University), Dehradun 248002, India

<sup>4</sup>Department of Electrical and Electronics Engineering, National Institute of Technology, Delhi 110040, India

<sup>5</sup>College of Computing, Khon Kaen University, Khon Kaen 40000, Thailand

<sup>6</sup>College of Electrical and Mechanical Engineering, Addis Ababa Science and Technology University, Addis Ababa 16417, Ethiopia

Correspondence should be addressed to Arfat Ahmad Khan; arfatkhan@kku.ac.th and Frie Ayalew; frie.ayalew@aastu.edu.et

Received 9 June 2022; Revised 6 July 2022; Accepted 30 July 2022; Published 1 September 2022

Academic Editor: Anastasios D. Doulamis

Copyright © 2022 Ashish Shukla et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The focus of this research is to isolating and identifying bacteria that produce calcite precipitate, as well as determining whether or not these bacteria are suitable for incorporation into concrete in order to enhance the material's strength and make the environment protection better. In order to survive the high "potential of hydrogen" of concrete, microbes that are going to be added to concrete need to be able to withstand alkali, and they also need to be able to develop endospores so that they can survive the mechanical forces that are going to be put on the concrete while it is being mixed. In order to precipitate  $\text{CaCO}_3$  in the form of calcite, they need to have a strong urease activity. Both *Bacillus sphaericus* and the *Streptococcus aureus* bacterial strains were evaluated for their ability to precipitate calcium carbonate ( $\text{CaCO}_3$ ). These strains were obtained from the Department of Biotechnology at GLA University in Mathura. This research aims to solve the issue of augmenting the tension and compression strengths of concrete by investigating possible solutions for environmentally friendly concrete. The sterile cultures of the microorganisms were mixed with water, which was one of the components of the concrete mixture, along with the nutrients in the appropriate proportions. After that, the blocks were molded, and then pond-cured for 7, 28, 56, 90, 120, 180, 270, and 365 days, respectively, before being evaluated for compressibility and tensile strength. An investigation into the effect that bacteria have on compression strength was carried out, and the outcomes of the tests showed that bacterial concrete specimens exhibited an increase in mechanical strength. When compared to regular concrete, the results showed a maximum increase of 16 percent in compressive strength and a maximum increase of 12 percent in split tensile strength. This study also found that both bacterial concrete containing 106, 107, and 108 cfu/ml concentrations made from *Bacillus sphaericus* and *Streptococcus aureus* bacteria gave better results than normal concrete. Both cluster analysis (CA) and regression analysis (RA) were utilized in this research project in order to measure and analyze mechanical strength.

## 1. Introduction

Concrete is thought of as a homogeneous substance since it is created by combining cement, coarse and fine aggregate, and water in a certain ratio. Concrete is a porous material and is sensitive to various assaults such as chloride,  $\text{CO}_2$ , sulfate, freeze and thaw cycles, and others because it is made up of voids that are referred to as pores [1, 2]. Since these

pores are typically associated with one other, concrete is a porous substance. Concrete has a design life of fifty years, but owing to these assaults, it deteriorates considerably more quickly than expected. The infrastructure is built of concrete [3–5]. Cementitious concrete is the most often used building material and is also one of the most essential substances used in construction business. Its annual production is around  $10 \text{ km}^3$ /per year, and it is one of the very important materials

in the construction industry. Cement is the sole component that is made, whereas the rest are naturally occurring and sourced from the area [6]. The manufacturing of 1 ton of cement results in the release of around 1 ton of carbon dioxide, and the building industry is responsible for approximately fifty percent of the world's total CO<sub>2</sub> emissions. Because of its adaptability, concrete is employed in the construction of a wide variety of structures, including bridges, large buildings, off-shore constructions, airports, sidewalks, railroad beds, and deep foundations, despite the fact that it is fragile and has a low resistance to stress [7–9]. A great number of concrete buildings are plagued by early deterioration issues such as carbonation and chloride attack, both of which ultimately result in the buildings needing to be repaired and retrofitted. Over the last several years, research into microbial (CaCO<sub>3</sub>) calcium carbonate has been more popular in the field of building engineering. It is seen as a potentially fruitful novel method for extending the useful life of cement-based buildings [10–12]. The CaCO<sub>3</sub> calcium carbonate precipitation that results from metabolic activity of various microbe species, such as sulfate-reducing microbes, ureolytic microbes, nitrate reducing microbes, and oxidation of organic microbes, is what allows this method to self-heal the gaps in the concrete and the inevitable microcracks that will form in the concrete. For instance, ureolytic microbes are responsible for the production of an enzyme known as urease, which decomposes urea into carbonate ions [13, 14]. According to the findings of our earlier studies, the process of carbonate precipitation by microbes was carried out by utilizing ureolytic microbes. Since they produce an enzyme called urease, these microbes are able to affect the precipitation of CaCO<sub>3</sub> in a given environment. This enzyme catalyzes the hydrolysis of urea into carbon dioxide and ammonia, which ultimately results in a rise in both the pH and the amount of carbonate in the surrounding/environment in which the microbes are found. The incorporation of microbial species into concrete has increased its strength and endurance, providing additional benefits in the form of environmentally friendly and cost-effective alternatives. Since bacteria thrive at alkaline pH, the concrete's resistance to alkali assault, chemical attacks, freeze-thaw strike, and drying shrinkage is significantly improved [15–17]. In a similar vein, the compression strength of Bio-Concrete was the subject of a great deal of research in order to evaluate the efficacy of bacteria-based self-healing in the existing body of scholarly work. This is due to the fact that the strength of concrete is regarded as an essential metric that represents the uniformity of a concrete mix as well as the components of the material being used [18–20]. The compressive strength test has a direct bearing on the general performance of the concrete as well as its attributes. As a result, the compression strength technique has been widely utilized to analyze the process of microbes-based self-healing in cementitious concrete, combining bacteria and associated chemical compounds in the research that has been published [21, 22]. An examination of the relevant literature reveals that virtually all bacteria are capable of producing calcium carbonate (CaCO<sub>3</sub>) as a precipitate; however, for the selected bacteria to provide the

greatest advantages, they must be alkaliphilic and thermophilic. *Bacillus sphaericus*, which is an alkaliphile and precipitates (CaCO<sub>3</sub>) calcium carbonate with a higher density, and *Streptococcus aureus*, which is also an alkaliphile and increases the compressive strength of concrete, are both beneficial microorganisms [23, 24]. For instance, the research group “Chattopadhyay et al.” found that the compression strength of bacterial mortar increased with time, specifically at seven, twenty-eight, and fifty-six days, in comparison to the normal mix. The precipitation of CaCO<sub>3</sub>, which filled the gaps and improved the texture of the concrete as a result, was thought to be responsible for the increase in compression strength that was observed. Additionally, “Ryu et al.” came to the conclusion that the highest increase in compression strength was observed at a microbial cell concentration of 30 × 10<sup>5</sup> cfu/mL, and also that the compressive strength declined once this value was exceeded [25]. In a similar vein, the “Manikandan et al.” study indicated that a rise in compressive strength occurred even at a high concentration level of 10<sup>8</sup> cells/mL. In the same vein, the greatest notable improvement in compression strength was seen when there were 10<sup>6</sup> cells per milliliter of solution [26]. Predictions of compressive strength are extremely valuable because of the significant savings in both money and time that they provide. Because of this fact, researchers have been motivated to construct a mathematical model that correctly forecasts the strength of various forms of concrete. Despite this, there still is not a prediction formula or code requirements for estimating the strength properties of microorganism concrete [27, 28].

In order to find a solution to the problem of increasing the tensile and compression strengths of concrete, the purpose of this study is to investigate potential options. Microorganisms grown in sterilized cultures were combined with water, which was one of the components of the concrete mixture, along with the nutrients in the right amounts. After that, the blocks were cast, and then they were pond-cured for 7, 28, 56, 90, 120, 180, 270, and 365 days, respectively, before being assessed for compressibility and tensile strength. In order to assess and investigate the nature of mechanical strength, this research study made use of both cluster analysis (CA) and regression analysis (RA). After a thorough study of old papers, the researcher found that not much research has been done with *Bacillus sphaericus* and *Streptococcus aureus* bacteria and also found that till date no researcher has done a thorough study of the results through cluster analysis for *Bacillus sphaericus* and *Streptococcus aureus* bacterial concrete. The researcher used three different concentrations, namely 10<sup>6</sup>, 10<sup>7</sup>, and 10<sup>8</sup> cfu/ml, of *Bacillus sphaericus* and *Streptococcus aureus* bacteria to determine the relationship between bacterial concrete and normal concrete.

This paper shows the introduction of the study in Section 1 and the materials used for the research in Section 2, bacteria (*Bacillus sphaericus* and *Streptococcus aureus*) used in Section 2.1, and research methodology in Section 3.1; Section 3.2 shows bacterial culture techniques, Section 3.3 shows concrete mix design, Section 4 shows results and discussion, and Section 5 shows conclusions.

## 2. Materials Used

OPC (ordinary Portland cement) of the grade 43 was utilized in this experimental endeavor. In accordance with IS 4031-1996 [29], this OPC was evaluated for both its physical qualities and its optical properties. The specific gravity, Blaine's fineness [30], soundness, and compressive strength of Portland cement were all measured and analyzed to identify their respective qualities [31, 32]. In this particular project, the fine aggregates consisted of river and crushed stone sand, both of which were sourced locally and were readily accessible. Sand from rivers typically has a specific gravity of about 2.68. As a coarse aggregate for this project, crushed granite-shattered stone with a nominal size of 20 millimeters is being employed [33–35]. In the experimental study, safe and potable water that was readily available in the local area was utilized for all combinations. Table 1 presents the variables studied and the procedure of bacterial concrete.

*2.1. Bacillus sphaericus Bacteria and Streptococcus aureus Bacteria.* Both *Bacillus sphaericus* and *Streptococcus aureus* were obtained from the Department of Biotechnology at GLA University in Mathura. The fact that the microorganisms *Bacillus sphaericus* and *Streptococcus aureus* are able to survive in the highly alkaline surroundings of concrete during the formation of  $\text{CaCO}_3$  crystals on concrete suggests that the presence of the microbes had no negative implication on the hydration of reaction, which forms a dense  $\text{CaCO}_3$  crystal in liquid medium [39, 61]. Ureolytic, Gram +ve, anaerobic, and round-spore-forming *Bacillus sphaericus* and *Streptococcus aureus* at a temperature of four degrees Celsius were grown on nutrient agar slants, and they were subcultured once every fifteen days on a medium that had been filtered and sterilized [59, 62].

## 3. Methodology

*3.1. Research Methodology.* The gathering of particular experimental data that contributes to a better understanding of microorganism concrete and the features it possesses is the primary aim of the current experimental research (durability and strength). In the current experimental inquiry, tests on the behavior of hardened and fresh characteristics of ordinary concrete grade and normal grade concrete without and with the inclusion of microbes have been carried out. These studies were carried out as part of the experimental study. Concrete in its hardened state is subjected to the appropriate laboratory tests, which enable the concrete strength, such as compression strength and tension strength. Research process methodology is shown in Figure 1.

The primary purpose of this experimental inquiry is to investigate the strength of normal concrete and bacterial concrete. The current work may be broken down into three distinct stages:

Phase 1: growth of bacteria and culturing technique of bacterial concrete

Phase 2: to study the compression strength and split tensile strength of bacterial concrete

Phase 3: regression analysis and cluster analysis

*3.2. Culturing Technique of Bacteria.* Both *Bacillus sphaericus* and *Streptococcus aureus* were cultivated in a medium that was designed for the improved generation of  $\text{CaCO}_3$ . This medium includes baking soda, ammonium chloride, urea, and  $\text{CaCl}_2$ , and it was dissolved in distilled water [36]. In order to determine the growth curve, the colonies were cultivated in batch culture aerobic endurance at 37 degrees Celsius and 100 revolutions per minute for a period of time [37, 41]. During this process, aliquots of the cells were taken out for optical density measurements and standard plate enumeration. We used a UV-vis 3000plus dbl spectrometer [63] to determine the initial concentration at 600 nm, also known as OD 600 (Department of Biotechnology, GLA University, Mathura).

*3.3. Mix Design of Concrete.* Concrete of the grade M30 was mixed in accordance with the standards set out in IS 10262-1982 [64]. For an exposure level of moderate and a water to cement ratio of 0.46, the amount of cement necessary to produce  $1 \text{ m}^3$  of concrete is 400 kg. Casting concrete into cubes of size 100 millimeters by 100 millimeters by 100 millimeters is carried out to determine the concrete's compressive strength, and casting concrete into cylinders of size 100 millimeters in diameter and 200 millimeters in length is carried out to determine the concrete's split tensile strength. Mix design is shown in Table 2.

## 4. Result and Discussion

*4.1. Compression Strength.* The evaluation of the effect that the incorporation of microorganisms into the mixture has on the compression strength and stiffness of cement concrete blocks is another factor that is taken into consideration [65–67]. The compression strength of the concrete was enhanced as a result of the influence of the bacterial isolates, as can be seen in Figures 2, 3, and Table 3. It is clear that the compression strength of microorganism concrete is significantly higher than that of the conventional concrete. The increase in compressive strength that is brought about by bacteria is presumably brought about by the deposition of calcium carbonate [67] on the surfaces of the microbe cell and the gaps inside the concrete, which plugs the holes that are present in the binder matrix [68, 69]. The filling of the gaps inside the concrete with microbiologically generated concrete mixes is the primary cause of the increase in compression strength [70]. The findings of the study indicated that the addition of bacteria to concrete led to an increase in compressive strength, which, in turn, would lead to an improvement in the concrete's overall performance.

*4.2. Split Tensile Strength.* In order to evaluate the tension strength of each mixture, three cylinders measuring 100 mm by 200 mm were casted for every combination. In order to create these cylinders, a mold made of cast iron and steel was

TABLE 1: Variables studied and procedures of bacterial concrete.

Materials used	Durability and mechanical test	Existing structure
Pseudomonas bacteria [36–39]	Compression strength [40–42]	Self-healing [16, 43–45]
Bacillus subtilis [2, 4, 34]	Tension strength [5, 46–48]	Resettlement of original structure [1, 16, 49, 50]
Megaterium bacteria [5, 51, 52]	Flexural strength [3], [28, 53, 54]	
Cereus and sphaericus bacteria [55–57]	Durability characteristics [12, 19, 58–60]	

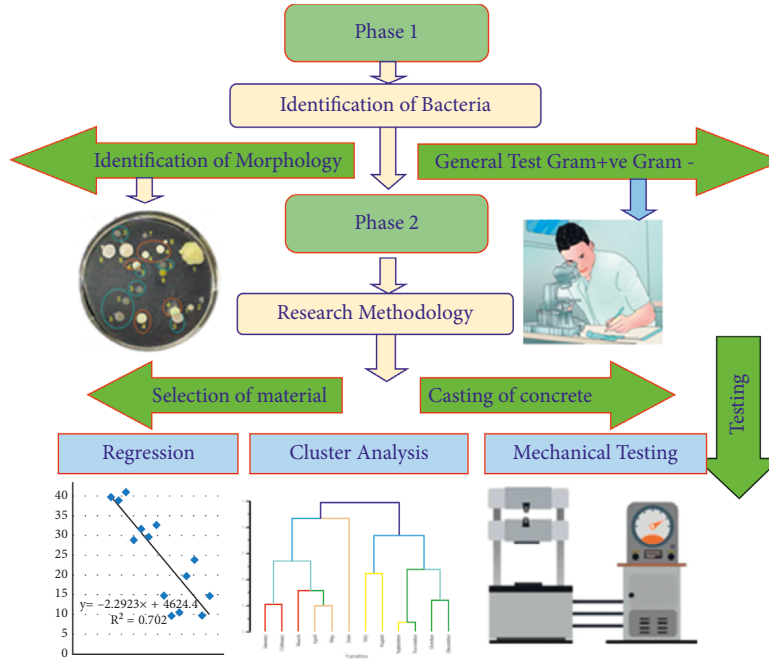


FIGURE 1: Research process methodology.

TABLE 2: Materials utilized in concrete mix design in detail.

Concrete	Mix Id	Sample	Bacteria cell concentration (ml)	OPC	Coarse aggregate	Fine aggregate	Water
Normal concrete	Conv.	A	—	400	1457	940	168
		B	—	400	1457	940	168
		C	—	400	1457	940	168
Bacillus sphaericus bacterial concrete	SP1	A	SP10 <sup>6</sup>	400	1457	940	168
		B	SP10 <sup>7</sup>	400	1457	940	168
		C	SP10 <sup>8</sup>	400	1457	940	168
	SP2	A	SP10 <sup>6</sup>	400	1457	940	168
		B	SP10 <sup>7</sup>	400	1457	940	168
		C	SP10 <sup>8</sup>	400	1457	940	168
	SP3	A	SP10 <sup>6</sup>	400	1457	940	168
		B	SP10 <sup>7</sup>	400	1457	940	168
		C	SP10 <sup>8</sup>	400	1457	940	168
Streptococcus aureus bacterial concrete	SA1	A	SA10 <sup>6</sup>	400	1457	940	168
		B	SA10 <sup>7</sup>	400	1457	940	168
		C	SA10 <sup>8</sup>	400	1457	940	168
	SA2	A	SA10 <sup>6</sup>	400	1457	940	168
		B	SA10 <sup>7</sup>	400	1457	940	168
		C	SA10 <sup>8</sup>	400	1457	940	168
	SA3	A	SA10 <sup>6</sup>	400	1457	940	168
		B	SA10 <sup>7</sup>	400	1457	940	168
		C	SA10 <sup>8</sup>	400	1457	940	168

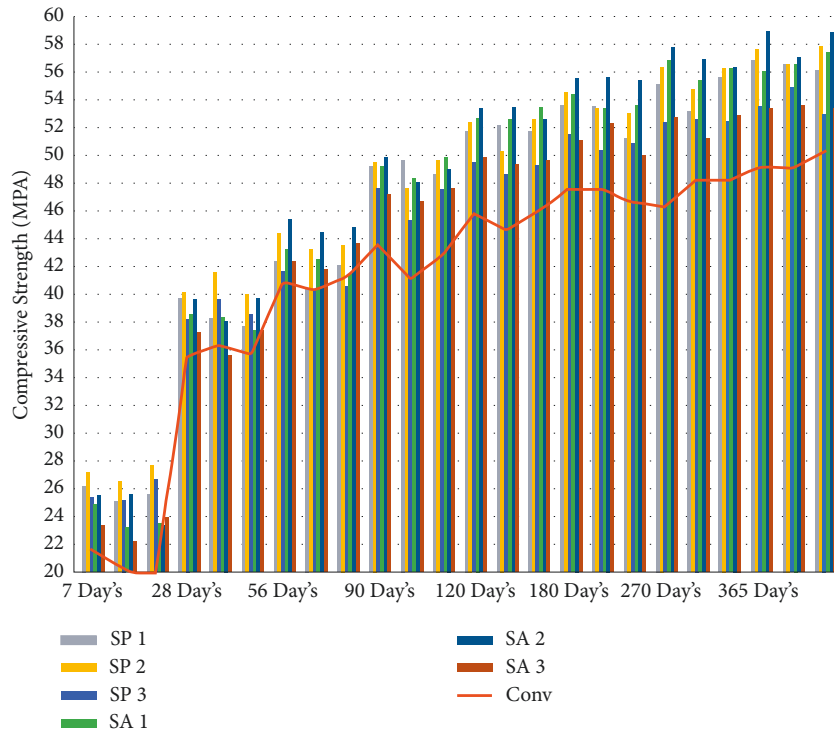


FIGURE 2: Graphical representation of compression strength result at different days' intervals.

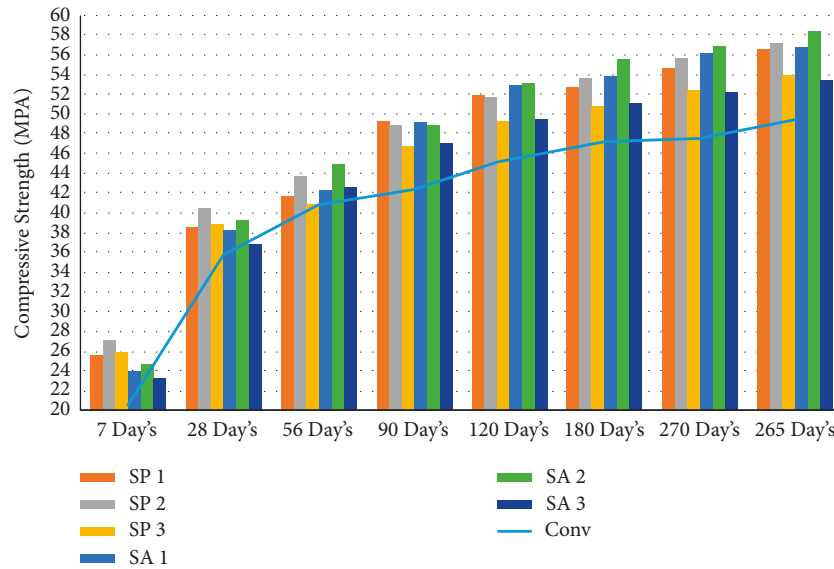


FIGURE 3: Graphical representation of average compression strength result at different days' intervals.

utilized. Casting and curing times for each mix percentage were as follows: seven days, twenty-eight days, fifty-six days, ninety days, one hundred eighty days, two hundred seventy days, and three hundred sixty-five days. In accordance with the Indian Standard code IS 516-1959 [71], a split tensile strength test is performed at the ages of seven days, twenty-eight days, fifty-six days, ninety days, one hundred twenty days, one hundred eighty days, two hundred seventy days, and three hundred sixty-five days using a compressive testing machine with a capacity of two thousand

kilonewtons. The loading part and the surface of the cylinder specimen are separated by a wooden strip so that there is no direct impact from the loading on the cylinder [55, 72, 73]. As shown in Figures 4, 5, and Table 4, the effect of the bacterial isolates resulted in an increase in the split tension strength of the concrete. This improvement can be attributed to the fact that the concrete was allowed to cure at a higher temperature [58]. It would suggest that the addition of *Bacillus sphaericus* and *Streptococcus aureus* was responsible for the rise in the splitting tension strength.

TABLE 3: Compression strength result at different days' intervals.

Day	Sample	Normal concrete	Bacillus sphaericus bacterial concrete			Streptococcus aureus bacterial concrete		
			10 <sup>6</sup> cfu/ml	10 <sup>7</sup> cfu/ml	10 <sup>8</sup> cfu/ml	10 <sup>6</sup> cfu/ml	10 <sup>7</sup> cfu/ml	10 <sup>8</sup> cfu/ml
7 days	A	21.71	26.32	27.31	25.53	24.91	25.61	23.36
	B	20.39	25.16	26.63	25.32	23.31	25.74	22.34
	C	19.86	25.72	27.81	26.72	23.67	23.31	23.91
28 days	A	35.61	39.72	40.19	38.32	38.61	39.65	37.31
	B	36.32	38.31	41.62	39.61	38.39	38.23	35.64
	C	35.68	37.69	40.03	38.59	37.42	39.69	37.65
56 days	A	40.88	42.39	44.39	41.67	43.23	45.37	42.37
	B	40.32	40.61	43.27	40.31	42.64	44.47	41.86
	C	41.23	42.13	43.53	40.63	41.35	44.83	43.73
90 days	A	43.61	49.24	49.57	47.67	49.23	49.91	47.21
	B	41.08	49.67	47.62	45.34	48.27	48.07	46.63
	C	42.71	48.63	49.67	47.57	49.86	49.03	47.67
120 days	A	45.83	51.71	52.37	49.54	52.71	53.43	49.87
	B	44.65	52.16	50.33	48.62	52.65	53.41	49.31
	C	45.79	51.75	52.64	49.37	53.53	52.62	49.64
180 days	A	47.64	53.62	54.45	51.47	54.43	55.53	51.07
	B	47.64	53.49	53.43	50.39	53.41	55.61	52.31
	C	46.59	51.35	53.06	50.87	53.63	55.39	50.03
270 days	A	46.23	55.12	56.39	52.39	56.86	57.74	52.75
	B	48.23	53.19	54.71	52.61	55.37	56.81	51.21
	C	48.17	55.64	56.23	52.39	56.26	56.37	52.86
365 days	A	49.13	56.87	57.61	53.67	56.07	58.96	53.39
	B	49.08	56.61	56.63	54.89	56.61	57.07	53.61
	C	50.23	56.12	57.81	52.93	57.37	58.83	53.37

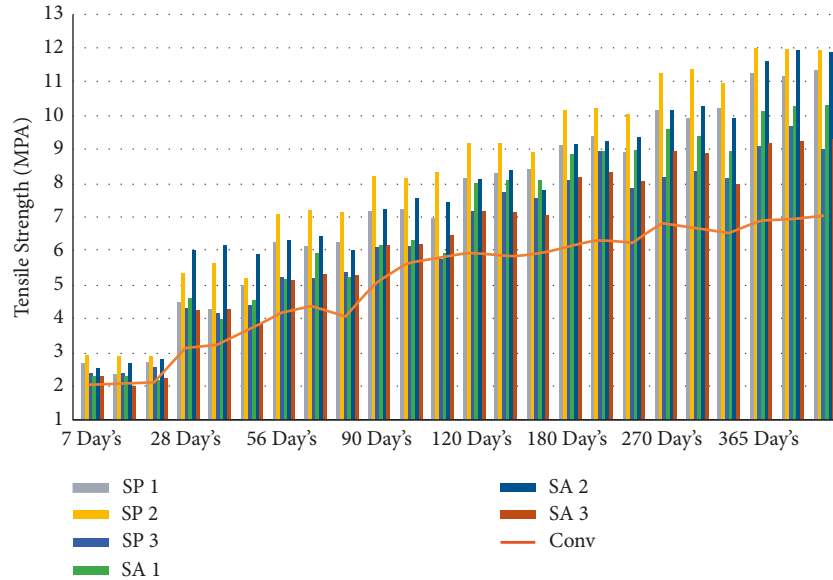


FIGURE 4: Graphical representation of split tension strength result at different days' intervals.

4.3. *Predicted Split Tension Strength.* The findings of the experiments were evaluated using regression analysis [43], which led to the discovery of the link between the compression strength of microbiological and normal concrete as well as tension strength of microbiological and normal concrete. This relationship may be expressed as equation (1), [74]. The predicted tensile strength's results at different day intervals are shown in Table 5 and Figures 6 and 7.

$$F_{\text{tensile}} = 0.23 f_{\text{compression}}^{0.73}. \quad (1)$$

4.4. *Regression Analysis.* The results of an experiment are depicted in Figure 8, which shows the link between the tension strength and the compression strength of bacterial concrete that was made with the bacteria *Bacillus*



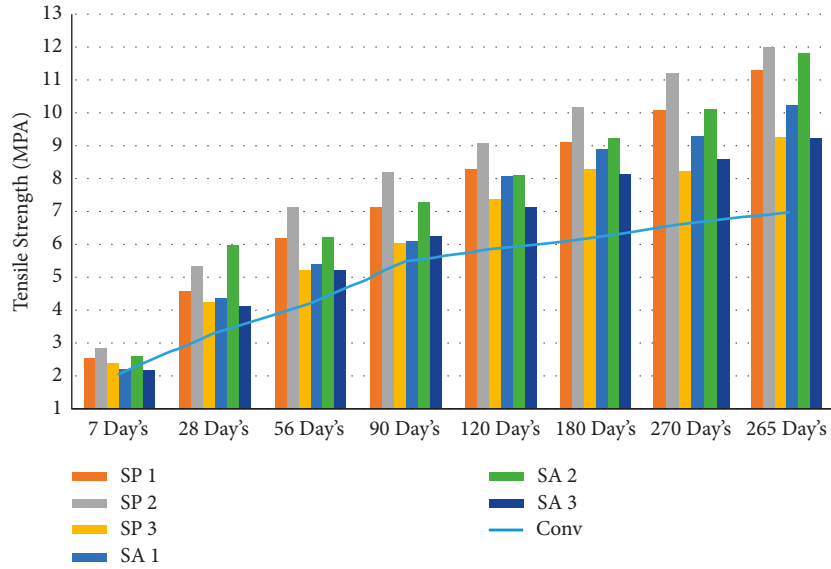


FIGURE 5: Graphical representation of average split tension strength result at different days' intervals.

TABLE 4: Tension strength result at different days' intervals.

Day	Sample	Normal concrete	Bacillus sphaericus bacterial concrete			Streptococcus aureus bacterial concrete		
			10 <sup>6</sup> cfu/ml	10 <sup>7</sup> cfu/ml	10 <sup>8</sup> cfu/ml	10 <sup>6</sup> cfu/ml	10 <sup>7</sup> cfu/ml	10 <sup>8</sup> cfu/ml
7 days	A	2.03	2.68	2.91	2.39	2.31	2.56	2.3
	B	2.11	2.39	2.89	2.41	2.29	2.63	2.16
	C	2.13	2.7	2.93	2.52	2.17	2.81	2.21
28 days	A	3.16	4.51	5.31	4.31	4.62	6.02	4.27
	B	3.23	4.31	5.62	4.16	3.98	6.13	4.31
	C	3.68	5.03	5.17	4.39	4.57	5.87	3.92
56 days	A	4.16	6.23	7.08	5.23	5.17	6.31	5.13
	B	4.39	6.13	7.16	5.16	5.93	6.42	5.31
	C	4.08	6.27	7.11	5.39	5.23	5.98	5.27
90 days	A	5.09	7.18	8.21	6.08	6.14	7.21	6.17
	B	5.67	7.23	8.18	6.17	6.31	7.52	6.19
	C	5.83	6.98	8.31	5.87	5.93	7.43	6.45
120 days	A	5.98	8.17	9.17	7.17	8.03	8.13	7.17
	B	5.87	8.31	9.19	7.77	8.12	8.39	7.13
	C	5.92	8.42	8.92	7.52	8.11	7.81	7.06
180 day's	A	6.13	9.16	10.16	8.13	8.87	9.16	8.16
	B	6.34	9.38	10.23	8.91	8.92	9.23	8.32
	C	6.24	8.93	10.08	7.86	8.99	9.33	8.05
270 days	A	6.87	10.19	11.26	8.17	9.62	10.19	8.91
	B	6.67	9.89	11.38	8.39	9.39	10.27	8.89
	C	6.54	10.21	10.97	8.17	8.91	9.89	7.98
365 days	A	6.91	11.31	12.03	9.16	10.11	11.61	9.17
	B	6.97	11.22	11.98	9.71	10.27	11.91	9.23
	C	7.08	11.34	11.91	8.98	10.33	11.89	9.31

sphaericus and Streptococcus aureus. The linear equation displays not only the percentage correlations between compression strength and split tensile strength ( $\sigma$ ) [75] but also the regression coefficients ( $R^2$ ) [76] that were derived from the equation that was shown below.

$$y = 0.249_x - 4.4595. \tag{2}$$

The following value for the regression coefficient,  $R^2$  (0.8813), indicates that the regression line as well as the statistics of compression strength and split tensile strength

TABLE 5: Predicted tensile strength result at different days' intervals.

Day	Sample	Normal concrete	Bacillus sphaericus bacterial concrete			Streptococcus aureus bacterial concrete		
			10 <sup>6</sup> cfu/ml	10 <sup>7</sup> cfu/ml	10 <sup>8</sup> cfu/ml	10 <sup>6</sup> cfu/ml	10 <sup>7</sup> cfu/ml	10 <sup>8</sup> cfu/ml
7 days	A	2.038227629	2.503436224	2.571832482	2.448358056	2.404809379	2.453956321	2.294631801
	B	2.077793632	2.422404145	2.524926774	2.433640017	2.291045402	2.463043444	2.221051093
	C	2.038227629	2.461646233	2.606120893	2.531153289	2.316821444	2.291045402	2.333946661
28 days	A	3.121586521	3.380684022	3.409839839	3.293278443	3.31145373	3.376333718	3.229685754
	B	3.166899629	3.292651047	3.497988291	3.373846899	3.297668971	3.287630294	3.123506065
	C	3.126064779	3.253665605	3.399924843	3.310201448	3.236634043	3.37881986	3.251144495
56 days	A	3.452476917	3.545113386	3.6664519	3.501055467	3.596259797	3.725366843	3.543892297
	B	3.417887928	3.435816163	3.598688612	3.417269093	3.560363902	3.671274354	3.512701689
	C	3.474030048	3.529227066	3.614461166	3.437051316	3.481408306	3.692946494	3.626576596
90 days	A	3.619309138	3.954758866	3.97408953	3.862307051	3.954172543	3.993969665	3.835064357
	B	3.464799065	3.979940439	3.859349338	3.723568456	3.897734599	3.885938698	3.800612473
	C	3.564629716	3.91893403	3.979940439	3.856390786	3.991048415	3.942439324	3.862307051
120 days	A	3.752902069	4.098616165	4.136738945	3.972333635	4.156327281	4.197696287	3.991632728
	B	3.682116322	4.124623116	4.018477076	3.918345731	4.152873001	4.196549189	3.958862225
	C	3.750510674	4.100930361	4.152297184	3.962378139	4.203430041	4.151145462	3.978185501
180 days	A	3.860532523	4.208587947	4.256045786	4.084720819	4.254904531	4.317506874	4.061522956
	B	3.860532523	4.201136887	4.197696287	4.021973619	4.196549189	4.322046643	4.13327862
	C	3.79823223	4.077766587	4.176456144	4.049905639	4.209160903	4.309558025	4.000977445
270 days	A	3.776785136	4.294212764	4.366217491	4.137892149	4.392753571	4.442279551	4.158629545
	B	3.895376476	4.18392344	4.270871825	4.150569557	4.308422033	4.389933406	4.069647769
	C	3.891838301	4.323748602	4.357170323	4.137892149	4.358867196	4.365086974	4.164958342
365 days	A	3.948307546	4.393317523	4.434976112	4.21145244	4.348116202	4.510604955	4.195401974
	B	2.503436224	4.378646039	4.379775259	4.281124862	4.378646039	4.404590959	4.208014962
	C	4.012647003	4.350946363	4.446210333	4.168983907	4.421481136	4.503342676	4.194254644

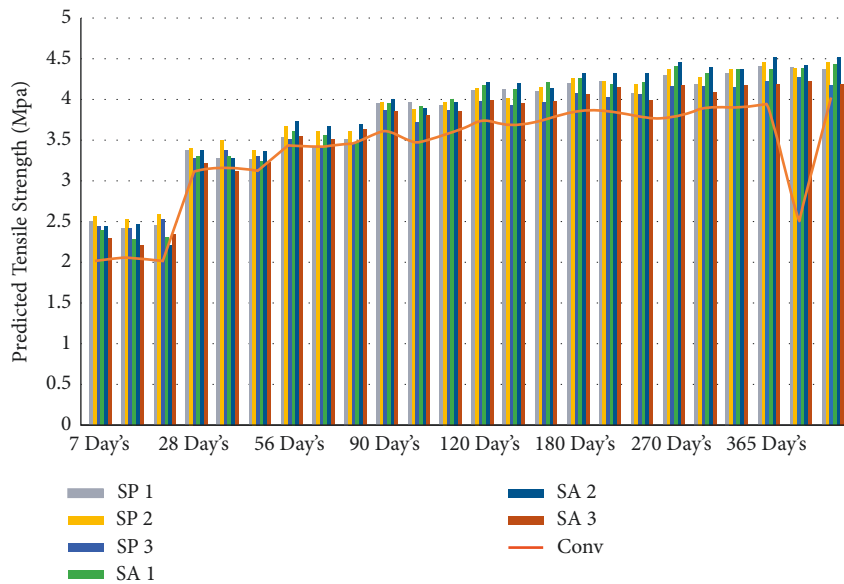


FIGURE 6: Graphical representation of predicted split tension strength result at different days' intervals.

values has a solid connection with one another. As can be seen from equation (2), the tensile strength improves as the compressive strength does.

4.5. Cluster Analysis. Multivariate statistical techniques are typically utilized for the categorization, analysis, and interpreting of big data sets. These approaches are also employed for the decrease of a dimension of complicated datasets with a minimal loss of original data [77]. Cluster

analysis is a method of unsupervised pattern classification that organizes the objects into the categories (clusters) on the basis of their commonalities within a category and their differences from other categories. The findings of CA lend a hand in data interpretation and point to patterns in the data. The square root of the average squared of the distinctions among corresponding values is computed in order to derive the distance from site in the Euclidean distance, which is one of the measurements that is utilized most frequently in order

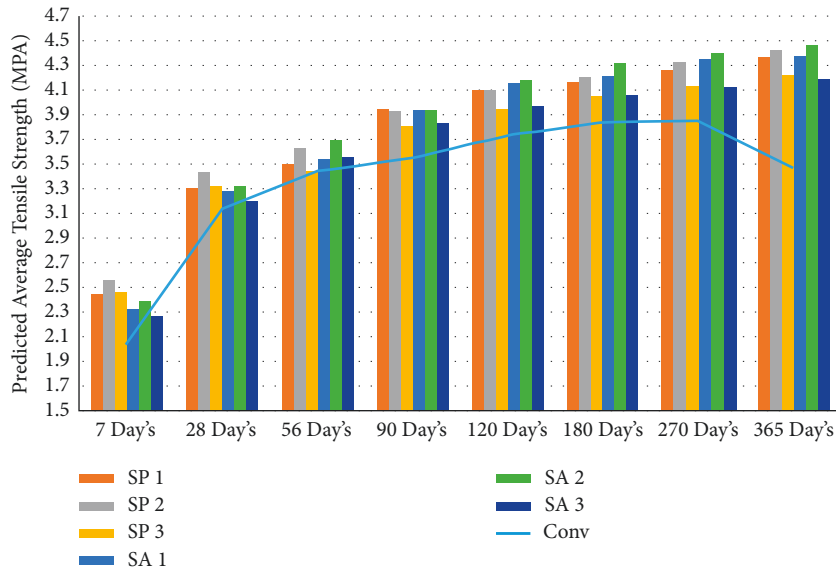


FIGURE 7: Graphical representation of average predicted split tensile strength result at different days' intervals.

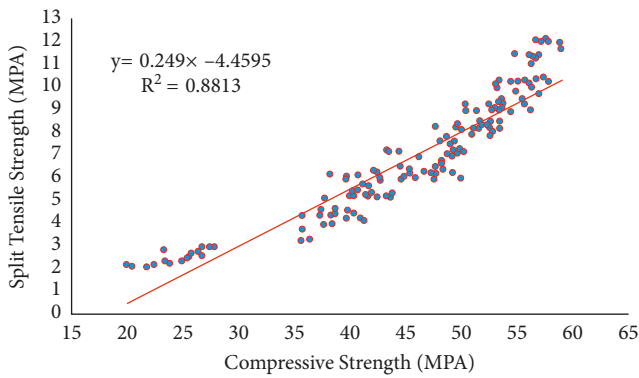


FIGURE 8: Linear regression analysis between compressive and tensile strength.

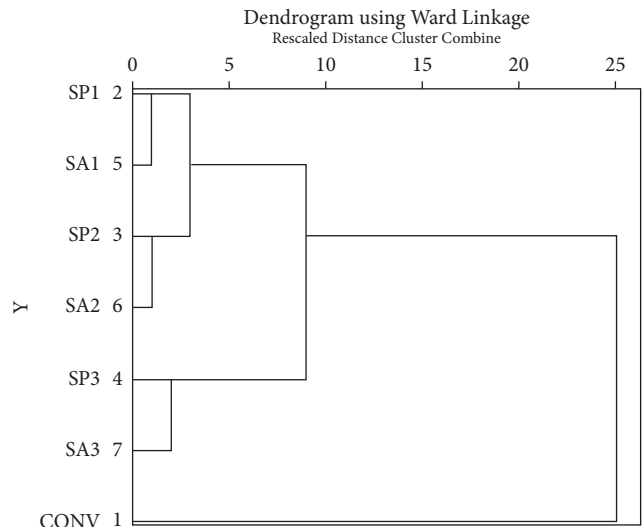


FIGURE 10: Cluster analysis result of split tensile strength.

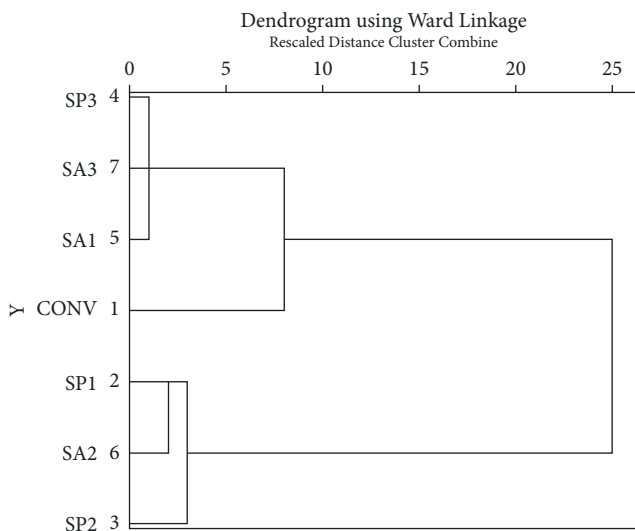


FIGURE 9: Cluster analysis result of compressive strength.

to determine the degree to which two cases are comparable to one another.

Through cluster analysis, it was found that the compression strength of bacterial concrete SP1, SA 1, SP 2, and SA 2 increases in the same way, and similarly, compression strength of SP3 and SA3 increases in the same group. Also, the compression strength of normal concrete sets itself apart from bacterial concrete. Cluster analysis results for compression strength are displayed in Figure 9.

Through cluster analysis, it was found that the tension strength of bacterial concrete SP 3, SA3, and SA1 increases in the same way, and tensile strength of SP 1, SA2, and SP 2 increases in the same group. Also, the split tensile strength of normal concrete sets itself apart from bacterial concrete. Cluster analysis results for split tensile strength are displayed in Figure 10.

## 5. Conclusion

- (i) The study of this research work has shown that the mechanical strength of concrete can be increased by using *Bacillus sphaericus* and *Streptococcus aureus* bacteria for bio-concrete.
- (ii) Using both bacteria in concrete with different concentrations showed that microbial concrete with a concentration of  $10^7$  cfu/ml produced better results for mechanical strength.
- (iii) This study also found that both bacterial concrete containing  $10^6$ ,  $10^7$ , and  $10^8$  cfu/ml concentrations made from *Bacillus sphaericus* and *Streptococcus aureus* bacteria gave better results than normal concrete.
- (iv) Value of the regression coefficient indicates that the regression line as well as the statistics of compression strength and split tensile strength values has a solid connection with one another.
- (v) Through cluster analysis, it was found that the compression strength of bacterial concrete SP1 and SA1 increases almost equally, similarly, the compression strength of bacterial concrete SP2 and SA2 increases almost equally, and both the compressive strength of both SP3 and SA increases almost equally in a similar group. And in the cluster analysis for split tensile strength, three groups were formed, in which in the first group the tensile strength of SP3, SA3, and SA1 increases almost equally, and in the second group the tensile strength of SP1, SA2, and SP2 is mutual, which grows almost evenly, and in the third group comes the normal concrete.
- (vi) Research organizations from all over the world-wide have been interested in the employment of microbes for the purpose of increasing the longevity of construction materials as a consequence of the encouraging findings obtained so far. Our grasp of the opportunities and constraints presented by biotechnological processes on construction materials may unquestionably benefit from the work that has been carried out by a number of research groups that have concentrated their attention on a variety of materials. Studies are still being conducted on the preservation of nutrients and microbial products since these factors have an effect on the survival, development, and creation of biofilms. "*Bacillus sphaericus* and *Streptococcus aureus* bacterial concrete" appears to become the most viable technique for generating crack-resistant concrete in the coming days, according to the findings of this research as well as the prior research that has been conducted.

## Data Availability

Data are available on request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## References

- [1] P. N. Reddy and B. V. Kavyateja, "Experimental study on strength parameters of self repairing concrete," *Annales de Chimie Science des Matériaux*, vol. 43, no. 5, pp. 305–310, 2019.
- [2] M. V. S. Rao, V. S. Reddy, and C. Sasikala, "Performance of microbial concrete developed using *Bacillus subtilis* JC3," *Journal of the Institution of Engineers*, vol. 98, no. 4, pp. 501–510, 2017.
- [3] N. N. T. Huynh, N. M. Phuong, N. P. A. Toan, and N. K. Son, "Bacillus subtilis HU58 immobilized in micropores of diatomite for using in self-healing concrete," *Procedia Engineering*, vol. 171, pp. 598–605, 2017.
- [4] B. Tayebani and D. Mostofinejad, "Penetrability, corrosion potential, and electrical resistivity of bacterial concrete," *Journal of Materials in Civil Engineering*, vol. 31, no. 3, Article ID 04019002, 2019.
- [5] C. Venkata Siva Rama Prasad and T. Vara Lakshmi, "Experimental investigation on bacterial concrete strength with *Bacillus subtilis* and crushed stone dust aggregate based on ultrasonic pulse velocity," *Materials Today Proceedings*, vol. 27, no. xxxx, pp. 1111–1117, 2020.
- [6] K. Vijay and M. Murmu, "Effect of calcium lactate on compressive strength and self-healing of cracks in microbial concrete," *Frontiers of Structural and Civil Engineering*, vol. 13, no. 3, pp. 515–525, 2019.
- [7] A. Shukla and N. Gupta, "Study on the efficacy of natural pozzolans in Cement Mortar," in *Calcined Clays for Sustainable Concrete*, Rilem Publications, pp. 469–480, Springer, Berlin, Germany, 2020.
- [8] A. Talaiekhazan, A. Keyvanfar, A. Shafaghat, R. Andalib, M. Z. A. Majid, and M. A. Fulazzaky, "A review of self-healing concrete research development," *J. Environ. Treat. Tech*, vol. 2, no. 1, pp. 1–11, 2014.
- [9] A. Shukla, N. Gupta, A. Gupta, R. Goel, and S. Kumar, "Study on the behaviour of green concrete by the use of industrial waste material: a review," *IOP Conference Series: Materials Science and Engineering*, vol. 804, no. 1, Article ID 012036, 2020.
- [10] V. Achal, A. Mukherjee, and M. S. Reddy, "Microbial concrete: way to enhance the durability of building structures," *Journal of Materials in Civil Engineering*, vol. 23, no. 6, pp. 730–734, 2011.
- [11] A. Gupta, N. Gupta, A. Shukla, R. Goyal, and S. Kumar, "Utilization of recycled aggregate, plastic, glass waste and coconut shells in concrete - a review," *IOP Conference Series: Materials Science and Engineering*, vol. 804, Article ID 012034, 2020.
- [12] V. Achal, A. Mukherjee, and M. Sudhakara Reddy, "Biogenic treatment improves the durability and remediates the cracks of concrete structures," *Construction and Building Materials*, vol. 48, pp. 1–5, 2013.
- [13] V. Achal, A. Mukherjee, and M. S. Reddy, "Effect of calcifying bacteria on permeation properties of concrete structures," *Journal of Industrial Microbiology & Biotechnology*, vol. 38, no. 9, pp. 1229–1234, 2011.
- [14] F. B. Silva, N. Boon, N. De Belie, and W. Verstraete, "Industrial application of biological self-healing concrete:

- challenges and economical feasibility,” *Journal of Commercial Biotechnology*, vol. 21, no. 1, pp. 31–38, 2015.
- [15] N. Gupta, A. Shukla, A. Gupta, R. Goel, and V. Singh, “A review on the selection of the variant water in concreting,” *IOP Conference Series: Materials Science and Engineering*, vol. 804, no. 1, p. 012037, 2020.
- [16] Y. Ç. Erşan, E. Gruyaert, G. Louis, C. Lors, N. De Belie, and N. Boon, “Self-protected nitrate reducing culture for intrinsic repair of concrete cracks,” *Frontiers in Microbiology*, vol. 6, p. 1228, 2015.
- [17] H. J. Kim, H. J. Eom, C. Park et al., “Calcium carbonate precipitation by *Bacillus* and *Sporosarcina* strains isolated from concrete and analysis of the bacterial community of concrete,” *Journal of Microbiology and Biotechnology*, vol. 26, no. 3, pp. 540–548, 2016.
- [18] J. Xu, Y. Tang, X. Wang, Z. Wang, and W. Yao, “Application of ureolysis-based microbial  $\text{CaCO}_3$  precipitation in self-healing of concrete and inhibition of reinforcement corrosion,” *Construction and Building Materials*, vol. 265, Article ID 120364, 2020.
- [19] M. J. C. Alonso, C. E. L. Ortiz, S. O. G. Perez et al., “Improved strength and durability of concrete through metabolic activity of ureolytic bacteria,” *Environmental Science and Pollution Research*, vol. 25, no. 22, Article ID 21451, 2017.
- [20] A. Shukla, N. Gupta, A. Gupta, R. Goel, and S. Kumar, “Natural pozzolans a comparative study: a review,” *IOP Conference Series: Materials Science and Engineering*, vol. 804, no. 1, Article ID 012040, 2020.
- [21] R. K. Verma, L. Chaurasia, V. Bisht, and M. Thakur, “Bio-mineralization and bacterial carbonate precipitation in mortar and concrete,” *Biosci. Bioeng.*, vol. 1, no. 1, pp. 5–11, 2015.
- [22] N. Chahal and R. Siddique, “Permeation properties of concrete made with fly ash and silica fume : influence of ureolytic bacteria,” *Construction and Building Materials*, vol. 49, pp. 161–174, 2013.
- [23] S. Jena and K. C. Panda, “Effect of bacteria on the properties of concrete - a review,” *UKIERI Concr. Congr. Concr. Glob. Build.*, 2019.
- [24] O. Öztürk, G. Yıldırım, Ü. S. Keskin, H. Siad, and M. Şahmaran, “Nano-tailored multi-functional cementitious composites,” *Composites Part B: Engineering*, vol. 182, Article ID 107670, 2020.
- [25] Y. Ryu, K. E. Lee, I. T. Cha, and W. Park, “Optimization of bacterial sporulation using economic nutrient for self-healing concrete,” *Journal of Microbiology*, vol. 58, no. 4, pp. 288–296, 2020.
- [26] A. T. Manikandan and A. Padmavathi, “An experimental investigation on improvement of concrete serviceability by using bacterial mineral precipitation,” *Int. J. Res. Sci. Innov.* 3, vol. 2, pp. 46–49, 2015.
- [27] A. Shukla, N. Gupta, and K. Kishore, “Experimental investigation on the effect of steel fiber embedded in marble dust based concrete,” *Materials Today Proceedings*, vol. 26, pp. 2938–2945, 2020.
- [28] V. Nagarajan, T. K. Prabhu, M. G. Shankar, and P. Jagadesh, “A study on the strength of the bacterial concrete embedded with *Bacillus megaterium*,” *Int. Res. J. Eng. Technol.*, vol. 4, no. 12, pp. 1784–1788, 2017.
- [29] IS 4031- Part I, *Method of Physical Tests for Hydraulic Cement: Determination of Fineness by Dry Sieving*, Bur. Indian Stand, New Delhi, 2005.
- [30] P. Sharma, N. Sharma, P. Singh, M. Verma, and H. S. Parihar, “Examine the effect of setting time and compressive strength of cement mortar paste using iminodiacetic acid,” *Materials Today Proceedings*, vol. 32, no. 4, pp. 878–881, 2020.
- [31] S. Krishnapriya and D. L. Venkatesh Babu, “Isolation and identification of bacteria to improve the strength of concrete,” *Elsevier GmbH.*, vol. 174, pp. 48–55, 2015.
- [32] N. De Belie, “Application of bacteria in concrete: a critical evaluation of the current status,” *RILEM Tech. Lett.*, vol. 1, p. 56, 2016.
- [33] V. Nežerka, Z. Slížková, P. Tesárek, T. Plachý, D. Frankeová, and V. Petráňová, “Comprehensive study on mechanical properties of lime-based pastes with additions of metakaolin and brick dust,” *Cement and Concrete Research*, vol. 64, pp. 17–29, 2014.
- [34] S. Joshi, S. Goyal, A. Mukherjee, and M. S. Reddy, “Microbial healing of cracks in concrete: a review,” *Journal of Industrial Microbiology and Biotechnology*, vol. 44, no. 11, pp. 1511–1525, 2017.
- [35] R. Bansal, N. K. Dhami, A. Mukherjee, and M. S. Reddy, “Biocalcification by halophilic bacteria for remediation of concrete structures in marine environment,” *Journal of Industrial Microbiology and Biotechnology*, vol. 43, no. 11, pp. 1497–1505, 2016.
- [36] B. Radha Kiranmaye, J. Ray Dutta, A. Kar, C. Parimi, and S. Raju, “Optimization of culture parameters of *Pseudomonas alcaligenes* for crack healing in concrete,” *Materials Today Proceedings*, vol. 28, pp. 713–716, 2020.
- [37] S. Heidari Nonakaran, M. Pazhouhandeh, A. Keyvani, F. Z. Abdollahipour, and A. Shirzad, “Isolation and identification of *Pseudomonas* azotoformans for induced calcite precipitation,” *World Journal of Microbiology and Biotechnology*, vol. 31, no. 12, pp. 1993–2001, 2015.
- [38] Y. Ç. Erşan, N. Boon, and N. De Belie, “Microbial Self-Healing concrete: Denitrification as an Enhanced and Environment-Friendly Approach,” in *Proceeding of the 5th Int. Conf. Self-Healing Mater*, Durham, NC, USA, September 2019.
- [39] Y. S. Lee, H. J. Kim, and W. Park, “Non-ureolytic calcium carbonate precipitation by *Lysinibacillus* sp. YS11 isolated from the rhizosphere of *Miscanthus sacchariflorus*,” *Journal of Microbiology*, vol. 55, no. 6, pp. 440–447, 2017.
- [40] P. Pachavannan, C. Hariharasudhan, M. Mohanasundram, and M. Anitha Bhavani, “Experimental analysis of self healing properties of bacterial concrete,” *Materials Today Proceedings*, vol. 33, pp. 3148–3154, 2020.
- [41] P. Kumar Jogi and T. V. S. Vara Lakshmi, “Self healing concrete based on different bacteria: a review,” *Materials Today Proceedings*, vol. 43, no. xxxx, pp. 1246–1252, 2021.
- [42] N. Karimi and D. Mostofinejad, “*Bacillus subtilis* bacteria used in fiber reinforced concrete and their effects on concrete penetrability,” *Construction and Building Materials*, vol. 230, Article ID 117051, 2020.
- [43] S. Farhadi and S. Ziadloo, “Self-healing microbial concrete-a review,” *Materials Science Forum*, vol. 990, pp. 8–12, 2020.
- [44] S. Bifathima, T. V. S. Vara Lakshmi, and B. N. Matcha, “Self healing concrete by adding *Bacillus megaterium* MTCC with glass & steel fibers,” *Civil and Environmental Engineering*, vol. 16, no. 1, pp. 184–197, 2020.
- [45] Y. S. Lee and W. Park, “Current challenges and future directions for bacterial self-healing concrete,” *Applied Microbiology and Biotechnology*, vol. 102, no. 7, pp. 3059–3070, 2018.
- [46] N. Iswarya, R. Adalarasan, V. Subathra Devi, and M. Madhan Kumar, “Experimental investigation on strength and durability of light weight bacterial concrete,” *Materials Today Proceedings*, vol. 22, pp. 2808–2813, 2020.

- [47] P. Jagannathan, K. S. Satya Narayanan, K. devi arunachalam, and S. kumar annamalai, "Studies on the mechanical properties of bacterial concrete with two bacterial species," *Materials Today Proceedings*, vol. 5, no. 2, pp. 8875–8879, 2018.
- [48] R. Pei, J. Liu, S. Wang, and M. Yang, "Use of bacterial cell walls to improve the mechanical performance of concrete," *Cement and Concrete Composites*, vol. 39, pp. 122–130, 2013.
- [49] M. S. Jafarnia, M. Khodadad Saryazdi, and S. M. Moshtaghioun, "Use of bacteria for repairing cracks and improving properties of concrete containing limestone powder and natural zeolite," *Construction and Building Materials*, vol. 242, Article ID 118059, 2020.
- [50] N. De Belie and J. Wang, "Bacteria-based repair and self-healing of concrete," *Journal of Sustainable Cement-Based Materials*, vol. 5, no. 1, pp. 35–56, 2015.
- [51] R. Andalib, M. Z. Abd Majid, M. W. Hussin et al., "Optimum concentration of Bacillus megaterium for strengthening structural concrete," *Construction and Building Materials*, vol. 118, pp. 180–193, 2016.
- [52] X. Sun, L. Miao, L. Wu, and R. Chen, "Improvement of biocementation at low temperature based on Bacillus megaterium," *Applied Microbiology and Biotechnology*, vol. 103, no. 17, pp. 7191–7202, 2019.
- [53] A. Toghroli, P. Mehrabi, M. Shariati, N. T. Trung, S. Jahandari, and H. Rasekh, "Evaluating the use of recycled concrete aggregate and pozzolanic additives in fiber-reinforced pervious concrete with industrial and recycled fibers," *Construction and Building Materials*, vol. 252, Article ID 118997, 2020.
- [54] A. K. Parande, B. R. Babu, K. Pandi, M. S. Karthikeyan, and N. Palaniswamy, "Environmental effects on concrete using Ordinary and Pozzolana Portland cement," *Construction and Building Materials*, vol. 25, no. 1, pp. 288–297, 2011.
- [55] J. Y. Wang, N. De Belie, and W. Verstraete, "Diatomaceous earth as a protective vehicle for bacteria applied for self-healing concrete," *Journal of Industrial Microbiology and Biotechnology*, vol. 39, no. 4, pp. 567–577, 2012.
- [56] S. Luhar, "A review paper on self healing concrete," *Journal of Civil Engineering Research*, vol. 5, no. 3, pp. 53–58, 2015.
- [57] W. De Muynck, K. Cox, N. D. Belie, and W. Verstraete, "Bacterial carbonate precipitation as an alternative surface treatment for concrete," *Construction and Building Materials*, vol. 22, no. 5, pp. 875–885, 2008.
- [58] R. Jakubovskis, A. Jankutė, J. Urbonavičius, and V. Grišniak, "Analysis of mechanical performance and durability of self-healing biological concrete," *Construction and Building Materials*, vol. 260, Article ID 119822, 2020.
- [59] E. Schlangen and S. Sangadji, "Addressing infrastructure durability and sustainability by self healing mechanisms - recent advances in self healing concrete and asphalt," *Procedia Engineering*, vol. 54, pp. 39–57, 2013.
- [60] F. Nosouhian, D. Mostofinejad, and H. Hasheminejad, "Concrete durability improvement in a sulfate environment using bacteria," *Journal of Materials in Civil Engineering*, vol. 28, no. 1, pp. 1–12, 2016.
- [61] A. K. Parashar and A. Gupta, "Investigation of the effect of bagasse ash, hooked steel fibers and glass fibers on the mechanical properties of concrete," *Materials Today Proceedings*, vol. 44, pp. 801–807, 2021.
- [62] J. Jayaprakash, "Advances in construction materials and systems," *Proc. Int. Conf. Chennai*, vol. 2, 2017.
- [63] N. Chahal, R. Siddique, and A. Rajor, "Influence of bacteria on the compressive strength, water absorption and rapid chloride permeability of concrete incorporating silica fume," *Construction and Building Materials*, vol. 37, pp. 645–651, 2012.
- [64] I. S. Is 10262, *Indian Concrete mix design guide lines*, vol. 1982, New Delhi, Bur. Indian Stand, Article ID 10262.
- [65] V. Karthekeyan R, D. R. Manju, and S. T, "Study on strength and durability properties of bacterial concrete," *International Journal of Advanced Research in Science, Communication and Technology*, pp. 13–20, 2022.
- [66] A. Paknahad, M. Goudarzi, N. W. Kucko, S. C. G. Leeuwenburgh, and L. J. Sluys, "Calcium phosphate cement reinforced with poly (vinyl alcohol) fibers: an experimental and numerical failure analysis," *Acta Biomaterialia*, vol. 119, pp. 458–471, 2021.
- [67] A. K. Parashar and A. Gupta, "Experimental study of the effect of bacillus megaterium bacteria on cement concrete," *IOP Conference Series: Materials Science and Engineering*, vol. 1116, no. 1, Article ID 012168, 2021.
- [68] H. Rong, G. Wei, G. Ma et al., "Influence of bacterial concentration on crack self-healing of cement-based materials," *Construction and Building Materials*, vol. 244, Article ID 118372, 2020.
- [69] Y. Ç. Erşan, E. Hernandez-Sanabria, N. Boon, and N. De Belie, "Enhanced crack closure performance of microbial mortar through nitrate reduction," *Cement and Concrete Composites*, vol. 70, pp. 159–170, 2016.
- [70] M. N. H. Khan, G. G. N. N. Amarakoon, S. Shimazaki, and S. Kawasaki, "Coral sand solidification test based on microbially induced carbonate precipitation using ureolytic bacteria," *Materials Transactions*, vol. 56, no. 10, pp. 1725–1732, 2015.
- [71] IS 516, *Method of Tests for Strength of Concrete*, 2018.
- [72] R. P. Santos, T. M. Ramos, B. M. Borges et al., "A selected bacterial strain for the self-healing process in cementitious specimens without cell immobilization steps," *Bioprocess and Biosystems Engineering*, vol. 44, no. 1, pp. 195–208, 2021.
- [73] A. R. Suleiman and M. L. Nehdi, "Effect of environmental exposure on autogenous self-healing of cracked cement-based materials," *Cement and Concrete Research*, vol. 111, no. May, pp. 197–208, 2018.
- [74] M. Sarkar, D. Adak, A. Tamang, B. Chattopadhyay, and S. Mandal, "Genetically-enriched microbe-facilitated self-healing concrete-a sustainable material for a new generation of construction technology," *RSC Advances*, vol. 5, no. 127, Article ID 105363, 2015.
- [75] N. Sharma and P. Sharma, "Effect of hydrophobic agent in cement and concrete: a Review," *IOP Conference Series: Materials Science and Engineering*, vol. 1116, no. 1, Article ID 012175, 2021.
- [76] H. M. Jonkers, A. Thijssen, G. Muyzer, O. Copuroglu, and E. Schlangen, "Application of bacteria as self-healing agent for the development of sustainable concrete," *Ecological Engineering*, vol. 36, pp. 230–235, 2010.
- [77] K. R. Singh, R. Dutta, A. S. Kalamdhad, and B. Kumar, "Risk characterization and surface water quality assessment of Manas River, Assam (India) with an emphasis on the TOPSIS method of multi-objective decision making," *Environmental Earth Sciences*, vol. 77, no. 23, p. 780, 2018.

## Research Article

# Detection of Subarachnoid Hemorrhage in Computed Tomography Using Association Rules Mining

Hathal Salamah Alwageed 

*College of Computer and Information Science, Jouf University, Sakakah, Saudi Arabia*

Correspondence should be addressed to Hathal Salamah Alwageed; [hswageed@ju.edu.sa](mailto:hswageed@ju.edu.sa)

Received 26 May 2022; Revised 13 July 2022; Accepted 22 July 2022; Published 31 August 2022

Academic Editor: Ateeq Ur Rehman

Copyright © 2022 Hathal Salamah Alwageed. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Subarachnoid hemorrhage (SAH) is one of the serious strokes of cerebrovascular accidents. There is an approx. 15% probability of spontaneous subarachnoid hemorrhage in all acute cerebrovascular accidents (CVAs). Most spontaneous subarachnoid hemorrhages are caused by ruptures of intracranial aneurysms, accounting for about 85% of all occurrences. About 15% of acute cerebrovascular disorders are caused by spontaneous subarachnoid hemorrhage. This illness is mostly caused by brain/spinal arteriovenous malformations, extracranial aneurysms, and hypertension. Computed tomography (CT) scan is the common diagnostic modality to evaluate SAH, but it is very difficult to identify the abnormality. Thus, automatic detection of SAH is required to recognize the early signs and symptoms of SAH and to provide appropriate therapeutic intervention and treatment. In this article, the gray-level cooccurrence matrix (GLCM) is used to extract useful features from CT images. Then, the New Association Classification Frequent Pattern (NCFP-growth) algorithm is applied, which is based on association rules. Then, it is compared with FP-growth methods with association rules and FP-growth methods without association rules. The experimental results indicate that the suggested approach outperforms in terms of classification accuracy. The proposed approach equates to a 95.2% accuracy rate compared to the conventional data mining algorithm.

## 1. Introduction and Background

SAH is a kind of acute cerebrovascular illness that can be caused by a variety of factors. Subarachnoid space can be filled with blood if a blood vessel ruptures in the brain or spinal cord. The most frequent cause is cerebral thrombosis. A subarachnoid hemorrhage can be classified into two groups according to the cause: spontaneous and traumatic. About 15% of acute cerebrovascular disorders are caused by spontaneous subarachnoid hemorrhage. This illness is most caused by brain/spinal arteriovenous malformations, extracranial aneurysms, and hypertension. This syndrome can be caused by a variety of conditions, such as Moyamoya disease, vasculitis of the brain, malignant tumors, homological abnormalities, meningitis, encephalitis, and anticoagulation therapy issues [1, 2]. The most common subarachnoid hemorrhage occurs when an intracranial aneurysm ruptures or a vascular abnormality rupture occurs, while the others are uncommon. Generalized

nonaneurysmal nonperimesencephalic subarachnoid hemorrhages (PNSHs) and atraumatic convex subarachnoid hemorrhages (cSAHs) are the most common types of spontaneous subarachnoid hemorrhages. Rinkei offered the following concept of PNSH in 1991: the front of the midbrain is where the central section of the subarachnoid hemorrhage occurs. The bleeding might be accompanied by an enlargement of the annular cistern base. The longitudinal fissure cistern's front half is not entirely filled. Neither a cerebral hemorrhage nor an expansion of the lateral fissure cistern is apparent [3]. Approximately 15% of all spontaneous subarachnoid hemorrhages, according to some estimates [4], are brought on by PNSH. sSAH occurs in the sulcus gyrus of the superficial cerebral cortex. The cerebellum's surrounding big sections, the anterior and posterior longitudinal fissures, or any subcutaneous cisterns are not typically affected by bleeding [5], and only 7.45 percent of the time [6] is the ventricle affected. As a result, early identification and treatment are crucial in

preventing and treating subarachnoid hemorrhage complications. The most common kind of spontaneous subarachnoid hemorrhage is intracranial aneurysmal subarachnoid hemorrhage, which accounts for around 85% of cases [7]. Intracranial aneurysm rupture is a gradual process. The mortality rate of the first hemorrhage might reach 40%, and the disability rate can reach 33% [8, 9]. The mortality incidence of rerupture is as high as 60–70% if it is not recognized and treated promptly, and the damage is significant [10]. More than 60 million people have aneurysms, and each year over 200,000 people have a ruptured and hemorrhaged brain aneurysm, posing a serious risk to their lives and health and substantial financial and psychological costs to society. Therefore, early detection can be achieved with appropriate methods.

Intracranial aneurysms must be extracted, and appropriate therapeutic therapies must be carried out. Since its inception, data mining has attracted a growing amount of interest. It works well with large, incomplete, and noisy practical applications. Data mining is used to analyze data that has hidden value. The data's potential information is discovered through induction, generalization, and reasoning. Simultaneously, the data are continually enhanced through the mining process, allowing the data to be completely gathered, interpreted, and utilized to its full potential [11, 12]. New technologies are influencing medicine and research in related sectors. We cannot fully utilize and benefit from data until we store and analyze massive amounts of data. Due to the inability of most modern database systems to process large quantities of data, these data cannot be used for disease analysis, diagnosis, and pathological research. Data expertise cannot predict how data will develop in the future. In medical data mining, doctors can gather precise information about illnesses and how they are treated, improving hospital management. Data mining and knowledge discovery are discussed in this article, as well as how data mining is used in medical applications. The fundamental theories, broad structure, and primary technologies and methodologies of data mining are all thoroughly explained, followed by an examination of the unique characteristics of medical data. And then mix the two to come up with a process model for medical data mining. Using association rule extraction and its application to medical image mining, photographs are merged using association rule extraction using the connected theories of association rules and the essential methods of association rule extraction. The texture features of subarachnoid hemorrhage in CT images are extracted using a gray-level symbiosis matrix in order to provide supplemental classification and diagnosis. Computed tomography (CT) scan is the common diagnostic modality to evaluate SAH, but it is very difficult to identify an abnormality. Thus, automatic detection of SAH is required to recognize the early signs and symptoms of SAH and to provide appropriate therapeutic intervention and treatment. In this article, the gray-level cooccurrence matrix (GLCM) is used to extract useful features from CT images. Then NCFP-growth algorithm is applied, which is based on association rules. Then, it is compared with FP-growth methods with association rules and FP-growth methods without association rules. The experimental results indicate that the suggested approach

outperforms in terms of classification accuracy. Further, the article is divided into the following sections: Section 2 describes the literature review, Section 3 is the proposed methodology, Section 4 is the experimental analysis, and Section 5 is the conclusion.

*1.1. Contributions.* Contributions to this research paper are as follows:

- (i) A machine learning technique for detecting subarachnoid hemorrhages in CT using association rule mining is proposed.
- (ii) The GLCM is used to extract useful features from CT images.
- (iii) The accuracy rate is improved up to 95.2%.

## 2. Literature Review

A perimesencephalic nonaneurysmal subarachnoid hemorrhage (PNSH) contributes to approximately 21%–68% of all spontaneous subarachnoid hemorrhages with negative DSA. Most patients with this disease are asymptomatic, have a small bleeding site, have a successful treatment process, and are found to have no complications, such as vasospasm or hydrocephalus, and recurrences are limited. Compared to aneurysmal subarachnoid hemorrhage, cavity hemorrhage has dramatically different consequences. It is considered separate from benign subarachnoid hemorrhage. The literature argues that for a condition with low morbidity, death, and disability rates, it is relevant for complications induced by examination procedures to be fewer than 0.5 percent. As a result, the technique of examination we use should be noninvasive and effective. Not only does it decrease the patient's exposure to the inspection procedure, but it also has greater sensitivity, allowing it to rule out aneurysms with a high death and disability rate. To effectively diagnose PNSH, it must also have a larger negative predictive value. In addition, another major subtype of subarachnoid hemorrhage, spontaneous localized subarachnoid hemorrhage, was found this year. Only a few cases of cSAH had been recorded before Spitzer et al. described 12 cases in 2005 [5]. Kumar et al. found that it occurs in roughly 7.45 percent of people [6]. In addition, there has been no extensive study of its incidence and gender differences. It has a different cause of bleeding than PNSH, but it has a good prognosis and is categorized as a benign subarachnoid hemorrhage. Aneurysms should be recognized from such subarachnoid hemorrhages in terms of diagnosis and therapy. Digital subtraction angiography (DSA) is used to diagnose aneurysms, assess their preoperative status, and determine other vascular imaging indicators [12, 16]. DSA, on the other hand, has a complication risk of around 1–2 percent, with about 0.5 percent experiencing irreversible neurological impairment. It can be fatal in extreme circumstances. Other invasive treatments, such as angiography, are available. The drug's drawbacks, such as radiation danger, long inspection times, and high costs, limit its widespread use in intracranial aneurysm screening and



follow-up observation [17, 18]. It is particularly inappropriate for exclusion screening and follow-up of benign SAH such as PNSH. As a result, noninvasive examination approaches are gaining popularity. In modern clinics, CT angiography has become widely used to detect intracranial aneurysms, treat endovascular problems, and perform surgical procedures instead of using only DSA technology. These noninvasive vascular imaging approaches allow for the accurate diagnosis of aneurysms while eliminating the risks associated with cerebral angiography [17, 18].

Manual CT image detection, on the other hand, has low accuracy and efficiency, and the technology of data mining is helping to alleviate this problem. Image mining is a young field that has just recently evolved. Among the areas involved are computer vision, image processing, image retrieval, data mining, machine learning, database, and artificial intelligence. Image mining is still in the exploratory research stage, even though these disciplines are quite established in their respective domains. Many researchers, both at home and abroad, have actively explored this discipline and made significant efforts in the following areas:

- (1) Celestial image mining: this technology uses scientists' precisely classified sky photographs as the training set to create a model for detecting galaxies, and it has been used to effectively discover volcanoes on Venus [19].
- (2) Satellite remote sensing image mining [20]: satellite imagery is being used more frequently in several fields to solve surface challenges through remote sensing. When moving targets are detected in remote sensing photographs and stored with their original images in a database, a whole host of additional information can be retrieved, including connections to the moving targets.
- (3) Spatial data mining [21]: using this tool, it is possible to understand geography, identify spatial correlations, and establish links between spatial data and nonspatial data. People are hoping to build a geographic data cube and mine spatial data using it. In this subject, geographical data association analysis is a prominent research issue, and various methods have been presented. An image data mining software prototype developed at NASA's Jet Propulsion Laboratory, "Diamond Eye," automatically extracts semantic information from a picture and determines the topography of craters. Satellite detection and analysis have been very useful [22].
- (4) Medical image mining [23]: image mining systems have benefited from the availability of many medical photos. The physician has always documented the diagnosis along with medical images. Visual qualities in medical imaging may be connected to diagnosis data in a variety of ways. Medical imaging has become a specialized field. Some research organizations, for example, investigate the gap between damaged brain tissues. Some research organizations utilize similar approaches for judging early breast

cancer [24, 25]. The diagnostic record's correlation between features and pathological characteristics can aid doctors in determining the tumor's location. Multimedia Miner is a prototype application developed by Simon Frase University in 1998. In the development of this approach, the DB-Miner relational database mining system and C-BIRD were used [26]. By utilizing multidimensional analysis technologies, a multimedia cube management system is capable of building multimedia data cubes for a variety of purposes, including summary knowledge for categorization and association rule knowledge for association purposes. MM-Associator is one of the modules, and it mostly mines image association rules. Image size, color, and image description are among the data linked by these rules [27]. The prototype system is made up of three functional modules: this module describes multimedia data attributes using several abstract layers while allowing users to scroll up and down to examine data on different levels. (2) You can search for association rules using image or video data with MM-Associator. (3) MM-Classifier classifies multimedia materials and provides explanations for each class, according to Chen et al. [28]. Machine learning-based algorithms play a vital role in lowering fatality rates and effectively managing bleeding. An IoT classification system based on support vector machines and feed-forward networks is presented. The machine learning-based tool can help specialists diagnose and manage brain hemorrhages by providing information on the kind of hemorrhage. Wang et al. [29] gave an overview of how deep learning algorithms may be used to identify and classify bleeding in CT images automatically. Using AI-based technologies to automate the diagnostic process would eventually lead to a more effective and more timely cure. To detect bleeding accurately, they use a deep learning algorithm based on CNNs.

At this time, image mining research is rather developed, and it can do different processing on medical pictures, allowing difficult-to-observe lesions to become clearer while also providing a degree of auxiliary diagnosis. Using this technology, hospitals are able to communicate better with each other, diagnose and treat illnesses more efficiently, and reduce their workload on equipment, creating an improvement in overall medical quality.

### 3. Methodology

*3.1. Association Rule's Overview.* The link between distinct things that exist in the same event is referred to as association rules. Agrawal et al. suggested the extraction of association rules for the first time in 1993. Since its development in the 1990s, it has evolved into one of the most important data mining methods. On the one hand, mining association rules can provide us with association linkages at multiple conceptual levels. The association rule mining

technique can be applied to obtain association rules for laws between the different levels of a hierarchical tree describing a domain-related idea. Alternatively, various types of data sets have unique association rules. The associations returned by association analysis are evaluated by two indicators: the level of support and the level of confidence represents both the level of interest the community has in the rule and its reliability. Finding the rules that have higher support and confidence than the threshold rules with the least support and confidence is the goal of association rule mining. However, there may be times when low-support limitations are required, such as during sickness surveillance. Association rule mining usually consists of two steps: finding data points that meet the minimum degree of support, followed by finding frequently occurring data points. A strong association rule is constructed by taking frequent data points with the lowest confidence level as the second step. The first phase of this mining strategy, which involves successfully locating frequent data points, is the most difficult because it has the biggest effect on the algorithm's performance. These two-step mining procedures are used by many famous algorithms, such as Apriori, DHP, and others, for mining association rules.

### 3.2. Classification of Association Rules

- (1) Boolean association rules and numerical association rules are classified according to the type of variables they are associated with. An association between a quantitative item or attribute is specified by a quantitative association rule. A Boolean association rule is composed of items that are either present or absent.
- (2) It is separated into single-dimensional association rules and multidimensional association rules based on the dimensionality of the data included in the rules. A one-dimensional association rule deals with some associations in a single attribute and has just one dimension for each item or attribute. A multidimensional association rule is one that deals with the relationship between two or more unique attributes and has two or more dimensions.
- (3) Each association rule can be categorized into a single-level rule and a multilevel rule based on the level of abstraction from the data included in the rule. There are no objects or characteristics in a rule collection whose association is based on a single abstraction level. Items or characteristics from distinct abstract levels are involved in the multilayer association rules.

**3.3. Association Rules Process.** "Association rule mining" is the process of removing rules from a transaction database that satisfy the minimal support and confidence conditions. A more basic mining strategy is to calculate all viable rules,  $a$  and  $b$ . This tactic, however, is obviously ineffective. A little data collection can provide hundreds of rules. More than 80% of the rules could be removed if the minimal

support level and the confidence level were decreased to 25% and 50%, respectively. As a result, the rules must be pruned first to increase mining efficiency. The calculation method for the rule's support shows that the rule's  $X \rightarrow Y$  support is solely dependent on the support count of the item set  $\{X \cup Y\}$ . A significant number of association rule extraction algorithms are divided into two stages: finding common data points and finding association rules. The former is used to identify which data points people are interested in (i.e., those that obtain more support than a pre-determined support threshold), also known as frequent data points. These rules are derived from frequently occurring sets of items with confidence levels greater than the threshold.

### 3.4. Algorithm for Association Rule Mining

**3.4.1. Apriori Algorithm.** The Apriori approach [30], developed by Professor Agrawal for analyzing shopping basket data, is the most basic and extensively used algorithm in the study of association rules. By scanning the database several times, the method obtains single-layer Boolean association rules that require frequent data points from the Latin Apriori, which means "from the beginning." Recursion, which is based on frequency set theory, is an important idea. Prior knowledge of often recurring data points gives the Apriori algorithm its name. It mines frequent data points using a circular hierarchical search. This loop generates  $(k + 1)$ -data points using  $k$ -data points. The most common 1-items set is incredibly  $T_1$ , and  $T_1$  is then used to mine  $T_2$ , and  $T_2$  is then used to mine  $T_3$ . This process must be repeated for each layer of mining, up to the point where no more frequent items can be mined.  $T_1$  is generated by the algorithm, which then searches the database  $D$ , removes a portion of the item set from the candidate set, and returns  $T_2$ . It then scans the database  $D$  again, generating a candidate set of  $T_3$  through  $T_2$ . This step is continued until there are no more data points to add to the list. The Apriori algorithm has a unique feature. In compliance with the definition, an itemset cannot be a frequent itemset if it does not meet the minimal support condition. A new itemset  $aI$  will be created from the itemset  $I$  if a new item is added to it.  $I$  cannot be a frequent itemset since the number of occurrences must be less than  $I$ . Its converse statement is that if an item set fails to reach the minimal support level, then all feasible supersets fail as well. Following this, we describe how  $T_k$  designs  $T_{k+1}$  and uses Apriori characteristics to identify frequent data points through the two procedures of joining and deleting.

- (1) Connection step: to mine  $T_{k+1}$ , two  $T_k$  itemsets can be joined to form a candidate set of  $T_{k+1}$ , which is designated as  $L_{k+1}$ . The two data points in  $T_k$  are  $t_1$  and  $t_2$ . If the data points  $t_1$  and  $t_2$  are identical except for the final and penultimate items, other items are identical; however, the penultimate item is not identical. As long as the entry in each  $T_k$ 's  $t_1$  and  $t_2$  is expected to be in lexicographical order, then  $T_k$ 's  $t_1$

and  $t_2$  can be connected, and this connection method can ensure the presence of all and non-duplicate candidate item sets.

- (2) Deletion step: The above connection principle generates the candidate set  $L_{k+1}$ , which is a superset of  $T_{k+1}$ . There are two parts to the deleting procedure here as well. First, according to the Apriori property, eliminate all previously discovered supersets of infrequent data points. Then scan the database for data points with support less than the minimal support criterion, and mark them as infrequent data points.
- (3) We may deduce from the deletion step that each item put in  $L_{k+1}$  has to be searched in the database before being added to  $T_{k+1}$ . The Apriori algorithm's verification step is its bottleneck. In the example above, the database would need to be searched ten times, resulting in a substantial I/O load, and here the Apriori algorithm would gain most of its benefits. Following the discovery of all frequent data points, association rules must be generated using frequent data points. Most algorithms, including Apriori, use a similar rule generating mechanism. All nonempty subsets  $S$  are constructed for each frequent itemset  $T$ . For each nonempty subset, if  $\sigma(T)/\sigma(S) \geq \text{min\_conf}$ , a strong association rule  $S \rightarrow (T - S)$  is generated.

The general form of the Apriori Algorithm [30] is as follows:

```

L1 = {large 1-data points}
For (k = 2; Lk-1 ≠ ∅; k++) do
  Ck = Apriori-gen (Lk-1)
  for (all transactions t ∈ Ct, do
    increment c.count
  end
  Lk = {c ∈ Ck | c.count ≥ minsup}
End
Solution = ∪k Lk

```

**3.4.2. DHP Algorithm.** By interlinking frequent data points of the previous layer, frequent data points of the next layer can be linked to the next layer's candidate sets. After obtaining the frequent item sets of this layer from the candidate set, the cycle continues. As the candidate set grows, the screening algorithm becomes less efficient since it requires one-to-one comparisons with the database to compute the level of support for each item and reduce the number of candidate sets to a minimum, thus decreasing the number of comparisons, so that algorithmic speed can be improved. Using DHP (Direct Hash Table Prulling), you can reduce the number of superfluous candidate sets for a mining association rule through the use of a hash table structure. When building  $(k+1)$ -data points from  $k$ -data points, create a hash bucket that will be used later to further filter the set of candidates. At the same time, the database

needs to regularly be updated with fresh data points. Although resources may be needed to update the hash bucket and a database, the DHP technique dramatically lowers the number of candidate sets, improves performance, and minimizes comparisons.

**3.4.3. Algorithm for Partition.** Apriori uses the Partition algorithm to generate rules from frequent data points. On the other hand, the Partition technique segments the database to reduce the cost of each database scan. The essential premise is that if an item set is common over the entire database, it must also be frequent across a subset. A frequent itemset mining operation is performed on each segment of the database in order to determine the frequent itemset for that segment. The second stage involves combining all the segmented frequent item sets into a large candidate set, then filtering and verifying the true frequent item set of the whole database by comparing the large candidate set to the entire database. Because the Partition method uses the divide-and-conquer strategy, the entire procedure only searches the database twice, considerably lowering I/O usage. However, sorting the database before mining is required to eliminate duplicate frequent data points between various segments, which restricts the implementation of the Partition method to some extent.

**3.4.4. Algorithm for FP-Growth.** The FP-growth method may immediately build frequent data points without going through the candidate item set stage. The FP-growth algorithm likewise employs the divide-and-conquer technique, but it does so in two stages: after compressing the entire database into an FP-tree while maintaining item information, a series of condition databases corresponding to frequently occurring items are generated for the compressed database. The FP-growth method is the most distinct from the other modified Apriori algorithms discussed. They have distinct classification standards for common data points. FP-growth classifies frequent items in decreasing order of support, whereas the Apriori method divides frequent items into 1-item sets, 2-item sets, ..., and  $k$ -item sets according to their size; item mining is done in increasing order of length, whereas item mining is done in decreasing order of length. After sorting, the transaction's frequent elements are added to the FP-tree, which is then mined recursively.

**3.4.5. NCFP-Growth Algorithm.** The shortcomings of mining association standards have been repeatedly mentioned since the topic was first raised. Several additional criteria have been proposed to improve the analysis of association rules and avoid the development of fake association rules. In this article, the NCFP-growth algorithm is proposed as an improved frequent pattern tree construction based on the NCFP-growth algorithm. By effectively incorporating interest degree weights into the system, the FP-growth algorithm reduces the number of redundant and inaccurate rules. Furthermore, this strategy effectively compresses the search space of the algorithm and reduces

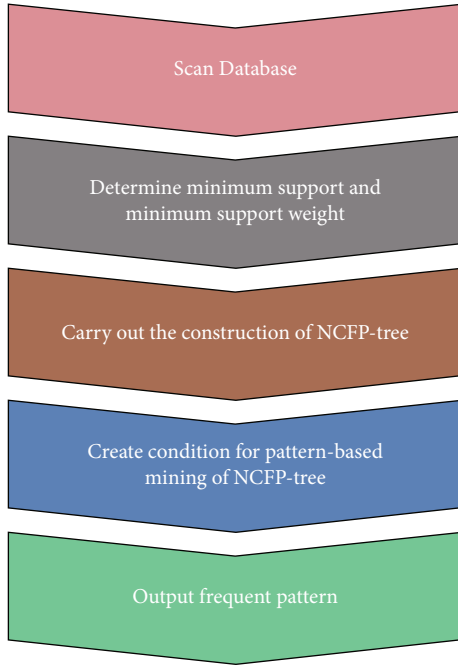


FIGURE 1: NCFP-growth algorithm.

the size of the tree and the system storage space compared to the FP-growth method. The NCFP-construction tree: the transaction database DB is the input; min sup is the minimal support threshold; min up is the minimum interest weight; min up is the whole collection of frequent patterns. This is how it's done: (1) examine the DB (transaction database); (2) identify the  $W$  elements with the highest degree of support using the minimal support minimum support; (3) arrange the  $W$  elements in table  $L$  in ascending order of degree of support. (4) Under NCFP-tree, create a root node and set its value to null. (5) Complete each transaction using the transaction database. (6) Sort the frequent things in each transaction that satisfy min up in the order in  $L$ ; the sorted table is labelled  $[p|P, T]$ , where  $p$  is the first element and  $P$  is the list of remaining elements. (7) If  $T$  has a child named  $N$  item name =  $p$  item name,  $N$ 's count is incremented by one; otherwise, a new node  $N$  with a count of one and a link to its parent node  $T$  is created. Utilizing the structure of a node chain, connect  $N$ 's node chain to the node with the same item name; if  $P$  is not empty, repeat steps 2 and 3 above. Afterwards, the NCFP-tree is built in the same manner as the FP-tree. As can be seen in Figure 1, the NCFP-growth algorithm is illustrated.

NCFP-growth algorithm shows that the system filters the original frequent items further when a new threshold is applied to them, thus reducing the possibility that too many absurd or redundant associations are generated, and enabling users to extract more practical association rules that are relevant to their needs.

**3.5. Gray-Level Cooccurrence Image's Matrix.** Each pixel in each image has a distinct or same gray level. The texture information in an image can be analyzed by considering the



FIGURE 2: Process model of CT image mining.

distance between two pixels. The conditional probability density function of approximated gray levels can be used to determine the gray levels of two pixels based on their cooccurrence matrix. To calculate the probability of another pixel with a gray level of  $j$  at a distance  $(D_x, D_y)$  having the same gray level as the pixel with a gray level of  $i$ , we use the following formula:

$$P(i, j, d, \theta) = \{ (x, y) f(x, y) = i, f(x + D_x, y + D_y) = j, x, y = 0, 1, 2, \dots, N - 1 \}. \quad (1)$$

There are four values:  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$ , and  $135^\circ$  for the gray level,  $x, y$  for the coordinates of the image pixels, and for the gray level. Thus, the gray-level pair  $(i, j)$  describes the texture information of the image. Clearly, the matrix resulting from the gray-level cooccurrence is symmetric. If, for example,  $M(1, 1) = 1$ , which means that there are only two horizontally adjacent gray level 1 pixels in the original image, then  $M(1, 1) = 1$ . Grayscale pixels with gray scales 1 and 2 are horizontally adjacent in the original image. Therefore,  $M(1, 2) = 2$ , meaning there are two grayscale pixels.

The image's coordinates are given by  $x, y$  image's  $\theta$  gray level as indicated by  $i, j$  and the image's direction is expressed by four values:  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$ , and  $135^\circ$ . The gray-level pair describes the texture information of the image in this way  $(i, j)$ . The GLCM that results is obviously symmetric. Using the GLCM element  $M(1, 1)$  as an example,  $M(1, 1) = 1$  indicates that in the original picture, only a pair of pixels with a gray level of 1 are horizontally contiguous. Because there are two pairs of pixels in the original image with gray scales of 1 and 2 horizontally contiguous,  $M(1, 2) = 2$ ,

**3.6. Association Rule Mining Method.** Thus, by converting gray-level cooccurrences from the arachnoid CT images into mathematical texture features, the process model of CT image mining is shown in Figure 2.

Image cropping, image noise reduction, and image enhancement are all required before extracting features. After the CT is complete, you can create a matrix of gray-level cooccurrences. The main goal of creating the matrix is to figure out what the gray level  $L$  and step are. Adding up the amount of data and computing the GLCM of the image determine two parameters of long  $D$ . Let's use the 288288 image format as an example. When the image is not compressed and step size  $D = 1$  and gray level  $L = 256$  is used, calculating the GLCM will take a very long time. In order to compress the original image,  $L$  was reduced and  $D$  was raised. With smaller gray levels  $L$  and larger step sizes  $D$ , the GLCM is more affordable to construct, but the less accurate the results, the more information lost. The higher the cost of building the gray-level cooccurrence matrix, the

TABLE 1: CT image feature data fragment.

PN	PD	$H$	$I$	$J$	$K$	$L$	$M$	Class
60	1	1	0	0	1	0	1	1
61	1	1	0	1	0	1	0	0
62	1	1	1	1	0	1	1	1
63	0	0	1	1	1	1	0	0
64	1	1	0	0	1	0	1	1
65	0	0	1	0	1	0	1	0
66	1	1	1	0	1	1	0	1
6667	1	1	1	1	0	1	0	1
68	1	0	1	1	1	1	0	1
69	0	1	0	0	0	1	1	0

higher the  $L$ , the smaller the  $D$ , and the more information kept. By identifying  $L$  and  $D$ , one can create the image's gray-level cooccurrence matrix. The matrix can be retrieved for its six attributes, which are energy, contrast, entropy, median, local stability, and correlation. A CT image mining database may be created by combining the following 6 elements of the arachnoid CT picture with the doctor's pre-diagnosis of the patient's subarachnoid hemorrhage (abbreviated as PD). PN, PD,  $H$ ,  $I$ ,  $J$ ,  $K$ ,  $L$ ,  $M$ , and Class are the attributes for each instance of each object stored in the database. PN indicates a probable diagnosis, PD indicates the physician's pre-diagnosis,  $H$  indicates the characteristics extracted from six gray-level cooccurrence matrices,  $K$ ,  $L$ , and  $M$  indicate whether the subarachnoid hemorrhage diagnosis is ultimately confirmed, and Class indicates the actual statement made by the physician. Using association rule mining, it is possible to diagnose and classify a database that has already been created. Here is a demonstration of how to mine a database section for subarachnoid hemorrhage diagnostic criteria. The present database is included in Table 1.

The following recommendations are based on clinical medical experience and intelligent image diagnosis principles; subarachnoid hemorrhage, or Class 1, is defined as a diagnosis that is based first on a clinical diagnosis and then on arachnoid CT image values that are at least 1, and at least two of the  $K$ ,  $L$ , and  $M$  values are 1. According to the study, the patient still suffered from subarachnoid hemorrhage, even though the initial diagnosis of No. 81 in the database did not match the initial criteria. This demonstrates that mining association guidelines should be used to assess the practicality and scientificity of diagnostic criteria.

## 4. Experiment and Analysis

**4.1. Data Collection.** There are several subarachnoid hemorrhage findings in CT images that appear to have irregular edges, dark surfaces, rough textures, and uneven grayscale distributions. Table 2 represents the Quantity and Grouping of Data image, which has 309 groups with a total of 642 images in each group. We assembled 50 sets of 1000 arachnoid CT images (both normal and pathological) from an affiliated hospital of a top-three medical hospital into an image database to explore how arachnoid CT images are applied clinically.

TABLE 2: Quantity and grouping of the data image.

Classification type	Group	Number of images
Normal	50	900
Abnormal	15	300

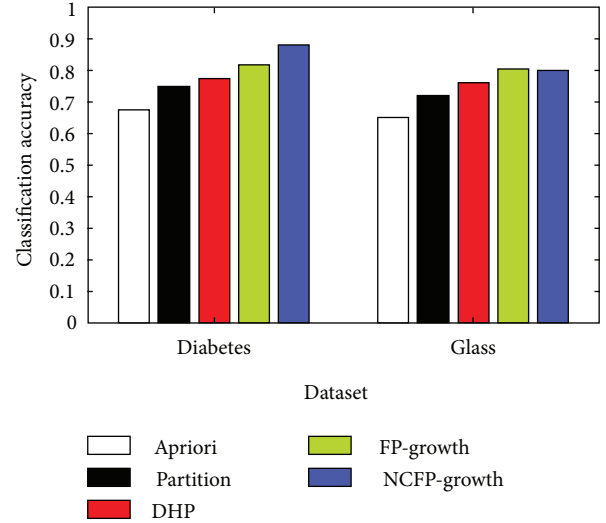


FIGURE 3: Comparison of five algorithms for classifying glass and diabetes.

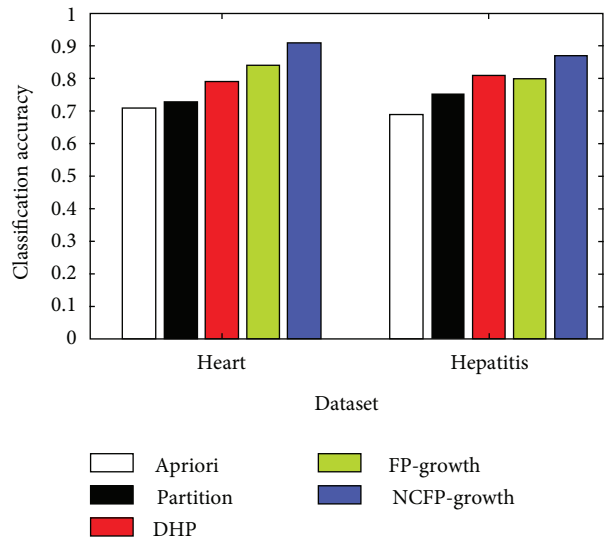


FIGURE 4: Hepatitis and heart disease: accuracy of five algorithms.

### 4.2. NCFP-Growth Compared to Other Growth Algorithms

**4.2.1. UCI Data Experiment Using Algorithms.** The NCFP-growth algorithm put forward in this study is among a number of algorithms that are assessed using common data from the data mining industry. In its ML Repository, UCI has eight datasets: diabetes, glass, hepatitis, iris, horses, labour, and led7. Figure 3 shows the comparison of five algorithms for classifying glass and diabetes, Figure 4 shows hepatitis and heart disease and the accuracy of five algorithms, Figure 5 shows the five algorithms for iris and horse

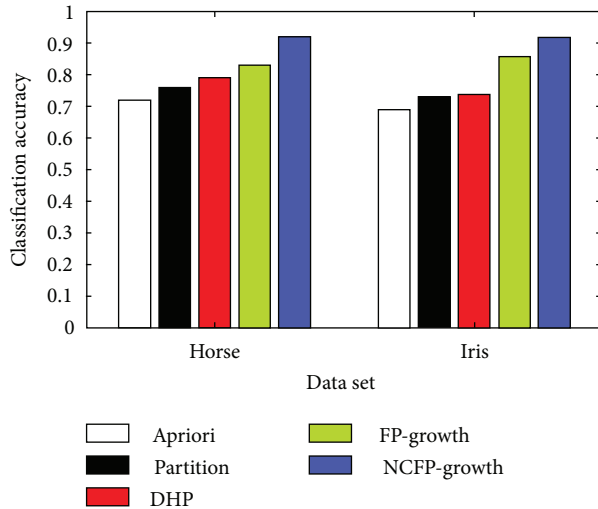


FIGURE 5: Five algorithms for iris and horse classification accuracy.

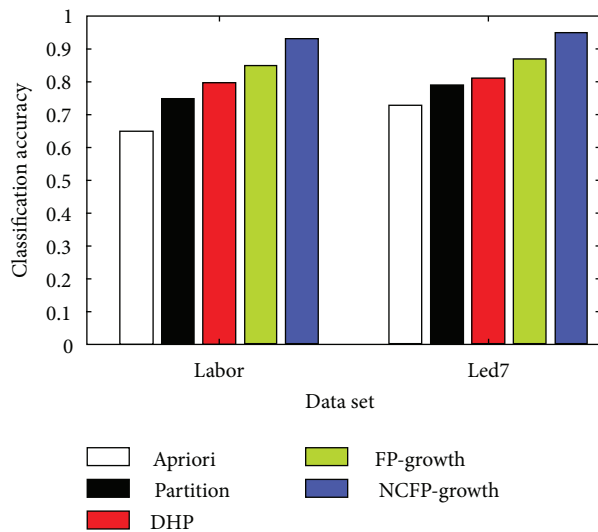


FIGURE 6: Five algorithms used to classify led7 and labour in accuracy.

classification accuracy, and Figure 6 shows the five algorithms used to classify led7 and labour in accuracy. We illustrate how to achieve comparable outcomes in comparison experiments using the five strategies outlined above.

The histogram clearly illustrates that the NCFP-growth algorithm outperforms the others in terms of accuracy. Due to the NCFP-growth algorithm, the mining accuracy is improved. Since it trims duplicate candidate's sets while searching for frequent item sets, the Partition algorithm's divide-and-conquer strategy can considerably increase the algorithm's performance.

*4.2.2. Analyzing the Subarachnoid Hemorrhage Dataset to Experiment with NCFP-Growth.* On the subarachnoid hemorrhage data set, we perform 10-fold cross-validation and compare the classification accuracy with those of the

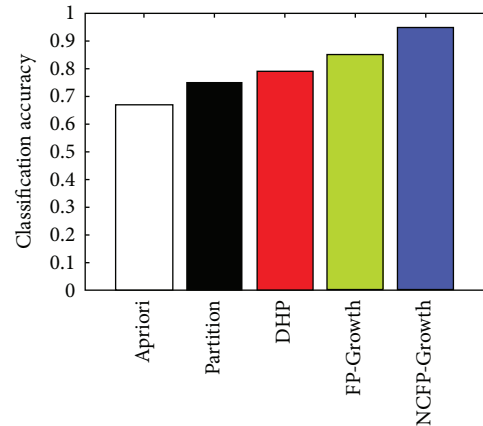


FIGURE 7: NCFP-growth classification accuracy and other algorithms.

TABLE 3: Number of FP-growth and association rules of NCFP-growth.

Algorithm	Number of CARs	Number of classifiers
FP-growth	8756	63
NCFP-growth	3122	26

different standard approaches discussed in the section, including Apriori, DHP, Partition, and FP-growth, which is depicted graphically in Figure 7. Experimental settings include 1% minimum support and 50% minimum constraint. Table 3 represents the number of FP-growth and association rules of NCFP-growth as follows.

Table 3 illustrates the fact that NCFP-growth is less complex than other algorithms, with 9252 candidate association rules instead of 3122 in FP-growth. NCFP-growth also reduced its classification rules from 63 to just 26 compared to FP-growth. Figure 7 illustrates that using the NCFP-growth algorithm equates to a 95.2% accuracy rate compared to the conventional data mining algorithm.

## 5. Conclusion

Data mining for medical purposes uses association rules extraction technologies to examine the application of these technologies. This article employs association rules mining to diagnose subarachnoid hemorrhage (SAH). The article explains in detail the approach used to implement this data mining technique. An Apriori algorithm for identifying frequent associations is presented, along with a review of the principles and fundamentals of association rule extraction. GLCM is the most powerful feature extraction used for characteristics of CT scans, followed by association rule image mining. Moreover, alternative mining methods based on different growth algorithms are compared to the NCFP-growth algorithm based on association rules. The NCFP-growth algorithm was found to be more accurate for SAH and it might be used to diagnose actual cases of such disease. The process segmentation is not applied in the proposed model, which is the main drawback of this article.

Additionally, the ensemble of deep learning models will be applied in the future for the diagnosis of SAH. The experimental results indicate that the suggested approach outperforms in terms of classification accuracy. The proposed approach equates to a 95.2% accuracy rate compared to the conventional data mining algorithm.

## Data Availability

The data used to support the findings of this study are included within the article.

## Conflicts of Interest

The author declares no conflicts of interest.

## Acknowledgments

The author carried out this study in the College of Computer and Information Science, Jouf University, Saudi Arabia., KSA.

## References

- [1] L. F. Zhang, J. Yang, Z. Hong et al., "Proportion of different subtypes of stroke in China," *Stroke*, vol. 34, no. 9, pp. 2091–2096, 2003.
- [2] Y. Q. Shi, *Practical Neurology*, Shanghai Science and Technology Press, Shanghai, China, 1994.
- [3] G. J. Rinkel, F. M. Wijdjcks, and M. Vermeulen, "Perimesencephalic nonaneurysmal subarachnoid hemorrhage: a follow-up study in 29 patients," *Surgical Neurology*, vol. 57, no. 3, p. 160, 2002.
- [4] S.R. A. SchwartzTH, "Perimesencephalic nonaneurysmal subarachnoid hemorrhage: review of the literature," *Neurosurgery*, vol. 39, no. 3, pp. 433–434, 1996.
- [5] C. Spitzer, M. Mull, V. Rohde, and C. M. Kosinski, "Non-traumatic cortical subarachnoid haemorrhage: diagnostic work-up and aetiological background," *Neuroradiology*, vol. 47, no. 7, pp. 525–531, 2005.
- [6] S. Kumar, R. P. Goddeau, M. H. Selim et al., "Atraumatic convexal subarachnoid hemorrhage: clinical Presentation, imaging Patterns, and etiologies," *Neurology*, vol. 74, no. 11, pp. 893–899, 2010.
- [7] R. Fogelholm, J. Hernesniemi, and M. Vapalahti, "Impact of early surgery on outcome after aneurysmal subarachnoid hemorrhage: a population-based study," *Stroke*, vol. 24, no. 11, pp. 1649–1654, 1993.
- [8] Y. Kato, H. Sano, K. Katada et al., "Application of three-dimensional CT angiography (3D-CTA) to cerebral aneurysms," *Surgical Neurology*, vol. 52, no. 2, pp. 113–122, 1999.
- [9] T. Inagawa and A. Hirano, "Ruptured intracranial aneurysms: an autopsy study of 133 patients," *Surgical Neurology*, vol. 33, no. 2, pp. 117–123, 1990.
- [10] J. B. Bederson, E. S. Connolly, H. H. Batjer et al., "Guidelines for the Management of Aneurysmal subarachnoid hemorrhage: a statement for healthcare professionals from a special writing group of the Stroke Council, American Heart Association," *Stroke*, vol. 40, pp. 994–1025, 2009.
- [11] F. I. H. Witten, "Data mining," *Practical Machine Learning Tools & Techniques with Java Implementations*, vol. 13, no. 1, p. 1, 2005.
- [12] K. J. Cios and G. William Moore, "Uniqueness of medical data mining," *Artificial Intelligence in Medicine*, vol. 26, no. 1–2, pp. 1–24, 2002.
- [13] J. Chen, K. Li, H. Rong, K. Bilal, N. Yang, and K. Li, "A disease diagnosis and treatment recommendation system based on big data mining and cloud computing," *Information Sciences*, vol. 435, pp. 124–149, 2018.
- [14] R. Anxionnat, S. Bracard, X. Ducrocq et al., "Intracranial aneurysms: clinical value of 3D digital subtraction angiography in the therapeutic decision and endovascular treatment," *Radiology*, vol. 218, no. 3, pp. 799–808, 2001.
- [15] H. J. Cloft, G. J. Joseph, and J. E. Dion, "Risk of cerebral angiography in patients with subarachnoid hemorrhage, cerebral aneurysm, and arteriovenous malformation: a meta-analysis," *Stroke*, vol. 30, no. 2, pp. 317–320, 1999.
- [16] R. A. Willinsky, S. M. Taylor, K. Terbrugge, R. I. Farb, G. Tomlinson, and W. Montanera, "Neurologic complications of cerebral angiography: prospective analysis of 2, 899 procedures and review of the literature," *Radiology*, vol. 227, no. 2, pp. 522–528, 2003.
- [17] P. Singh, S. Singh, and G. S. Pandi-Jain, "Effective heart disease prediction system using data mining techniques," *International Journal of Nanomedicine*, vol. 13, p. 121, 2018.
- [18] H. E. Westerlaan, A. M. van der Vliet, J. M. Hew, J. D. M. Metzemaekers, J. J. A. Mooij, and M. Oudkerk, "Magnetic resonance angiography in the selection of patients suitable for neurosurgical intervention of ruptured intracranial aneurysms," *Neuroradiology*, vol. 46, no. 11, pp. 867–875, 2004.
- [19] C. Ordonez and E. Omiecinski, "Discovering association rules based on image content," *Proceedings of the IEEE Advances in Digital Libraries Conference (ADL'990)*, Baltimore, MD, USA, May 1999.
- [20] M. Datcu and K. Seidel, "Image information mining: exploration of image content in large archives," *Proceedings of the IEEE Aerospace Conference Proceedings*, Big Sky, MT, USA, March 2000.
- [21] Ji Zhang, W. Hsu, and M. LiLee, "An information-driven framework for image mining," *Proceedings of the 12th International Conference on Database and Expert Systems Applications (DEXA)*, pp. 232–242, Munich, Germany, September 2001.
- [22] M. C. Burl, F. Charless, and R. Joseph, "Mining for image content," in *Systemics, Cybernetics, and Informatics/Information Systems: Analysis and Synthesis*, Orlando, FL, 1999.
- [23] D. Xu and F. Li, "Research and Application of CT Image Mining Based on Rough Sets Theory and Association rules [C]," *Proceedings of the IEEE International Conference on Computer Science & Information Technology*, pp. 392–394, IEEE, Chengdu, China, July 2010.
- [24] B. K. Velthuis, G. J. Rinkel, L. M. Ramos et al., "Subarachnoid hemorrhage: aneurysm detection and preoperative evaluation with CT angiography," *Radiology*, vol. 208, no. 2, pp. 423–430, 1998.
- [25] C. Glaser, R. Burgkart, A. Kutschera, K. H. Englmeier, M. Reiser, and F. Eckstein, "Femoro-tibial cartilage metrics from coronal MR image data: technique, test-retest reproducibility, and findings in osteoarthritis," *Magnetic Resonance in Medicine*, vol. 50, no. 6, pp. 1229–1236, 2003.
- [26] H. Laurent, "Wind extraction from meteorological water vapor channel image data," *Journal of Applied Meteorology*, vol. 32, no. 6, pp. 1124–1133, 1993.
- [27] I. Dnvsls, "Count of equivalent pixel (cep): a novel algorithm to extract most relevant images from the medical image

- database[J],” *European Journal of Translational and Clinical Medicine*, vol. 7, no. 3, pp. 5047–5055, 2020.
- [28] H. Chen, S. Khan, B. Kou, S. Nazir, W. Liu, and A. Hussain, “A Smart Machine Learning Model for the Detection of Brain Hemorrhage Diagnosis Based Internet of Things in Smart Cities,” *Complexity*, vol. 2020, Article ID 3047869, 10 pages, 2020.
- [29] X. Wang, T. Shen, S. Yang et al., “A deep learning algorithm for automatic detection and classification of acute intracranial hemorrhages in head CT scans,” *NeuroImage: Clinica*, vol. 32, Article ID 102785, 2021.
- [30] R. Agrawal and R. Srikant, “Fast algorithms for mining association rules,” in *Proceedings of the 20th International Conference on Very Large Data Bases* vol. 1215, VLDB edition, 1994.



## *Retraction*

# **Retracted: An AI-Driven Hybrid Framework for Intrusion Detection in IoT-Enabled E-Health**

### **Computational Intelligence and Neuroscience**

Received 28 November 2023; Accepted 28 November 2023; Published 29 November 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### **References**

- [1] F. Wahab, Y. Zhao, D. Javeed et al., “An AI-Driven Hybrid Framework for Intrusion Detection in IoT-Enabled E-Health,” *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 6096289, 11 pages, 2022.

## Research Article

# An AI-Driven Hybrid Framework for Intrusion Detection in IoT-Enabled E-Health

Fazal Wahab <sup>1</sup>, Yuhai Zhao,<sup>1</sup> Danish Javeed <sup>2</sup>, Mosleh Hmoud Al-Adhaileh,<sup>3</sup> Shahab Ahmad Almaaytah,<sup>4</sup> Wasiat Khan,<sup>5</sup> Muhammad Shahid Saeed <sup>6</sup>, and Rajeev Kumar Shah <sup>7</sup>

<sup>1</sup>College of Computer Science and Technology, Northeastern University, Shenyang 110169, China

<sup>2</sup>Software College, Northeastern University, Shenyang 110169, China

<sup>3</sup>Deanship of E-Learning and Distance Education, King Faisal University, P.O. Box 400, Al-Ahsa, Saudi Arabia

<sup>4</sup>Applied College in Abqaa, King Faisal University, Al-Ahsa, Saudi Arabia

<sup>5</sup>Department of Software Engineering, University of Science and Technology Bannu, Bannu, Pakistan

<sup>6</sup>Dalian University of Technology, Dalian 116024, China

<sup>7</sup>Sunway International Business School, Kathmandu, Nepal

Correspondence should be addressed to Rajeev Kumar Shah; [drrajeev@sunway.edu.np](mailto:drrajeev@sunway.edu.np)

Received 26 May 2022; Revised 22 July 2022; Accepted 26 July 2022; Published 21 August 2022

Academic Editor: Deepika Koundal

Copyright © 2022 Fazal Wahab et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

E-health has grown into a billion-dollar industry in the last decade. Its device's high throughput makes it an obvious target for cyberattacks, and these environments desperately need protection. In this scientific study, we presented an artificial intelligence (AI)-driven software-defined networking (SDN)-enabled intrusion detection system (IDS) to address increasing cyber threats in the E-health and internet of medical things (IoMT) environments. AI's success in various fields, including big data and intrusion detection systems, has prompted us to develop a flexible and cost-effective approach to protect such critical environments from cyberattacks. We present a hybrid model consisting of long short-term memory (LSTM) and gated recurrent unit (GRU). The proposed model was thoroughly evaluated using the publicly available CICDDoS2019 dataset and conventional evaluation measures. Furthermore, for proper validation, the proposed framework is compared with relevant classifiers, such as cu-GRU+DNN and cu-BLSTM. We have further compared the proposed model with existing literature to prove its efficacy. Lastly, 10-fold cross-validation is also used to verify that our results are unbiased. The proposed approach has bypassed the current literature with extraordinary performance ramifications such as 99.01% accuracy, 99.04% precision, 98.80 percent recall, and 99.12% F1-score.

## 1. Introduction

The internet of things (IoT) has been identified as an essential research domain for the present and coming decade. The applications of IoT have been integrated into industries and health areas to aid the people and emerged as industrial internet of things (IIoT) and IoMT. The IIoT revolution is exploding, resulting in massive monetary gains and automation [1]. On the other hand, the IoMT has also grown into a multibillion-dollar industry. While providing significant benefits, the pervasive and open nature of the IoMT ecosystem makes it a possible target for various emerging cyber

threats and attacks [2–5]. The extensive connectivity and continuous sharing of data of these devices make them a prime target of different threat actors that can execute anomalous activities against them [6]. The exploit's motivations are to obtain important information, steal money, and damage the system's resources [7–9]. As the number of linked IoT devices grows, critical infrastructure and assets of different organizations are also becoming vulnerable to numerous cyberattacks. Cyber threats could cost up to \$ 90 trillion by 2030 if no reasonable alternative is given before then [10, 11]. IoMT environments pose three issues as follows: The first is the heterogeneous network and dynamic

nature, the second is its hugely scattered design, and the last is the protocols that the IoT use to address concerns like computing limits and power consumption in network sensors [12, 13]. The most common issue in IoMT setups is keylogging, botnet attacks, and zero-day exploits [14–16].

The intruder’s primary purpose is to contaminate sensitive machines with different techniques, including denial-of-service (DoS) attacks, distributed denial of service (DDoS), and advanced persistent threats (APTs), in order to gain control and change their functioning [17, 18]. The nuclear program of Iran, for example, was targeted by the Stuxnet worm in 2010. Later, in 2013, Iranian hackers gained access to the dam’s ICS. In Ukraine, Black Energy malware caused a power outage for 230,000 people in 2015 [19]. As a result, these incidents demonstrated that typical cybersecurity methods, such as authentication, security rules, security firewalls, both software and hardware-based, and IDS, are no longer beneficial.

Similarly, the IIoT’s digital landscape is vulnerable to sophisticated hacking techniques, physical security risks, and a wide range of devices that can be easily infected by botnet attacks [20]. Furthermore, the IoMT demands a different detection mechanism for its environments due to low latency and resource limitations. Hence, such environments need a scalable, cost-effective, and adaptive intrusion detection mechanism against emerging cyber threats. The proposed network model is shown in Figure 1.

*1.1. Contribution.* The main contributions of this research are as follows:

- (i) We presented a novel, i.e., Cu-LSTM+ GRU SDN-enabled intelligent framework to detect threats quickly and effectively in the IoMT environment. The proposed SDN-enabled model does not overburden the IoMT resource.
- (ii) We employed a publicly available, state-of-the-art CICDDoS2019 dataset to evaluate the performance of the proposed model.
- (iii) We evaluated the proposed model’s performance by employing two existing benchmark algorithms, i.e., Cu-GRU-DNN and Cu-BLSTM, which were trained and assessed on the same dataset.
- (iv) To comprehensively assess the proposed model’s performance, we have compared it to the existing literature.
- (v) For a better assessment, we have utilized the standard evaluation metrics.
- (vi) Finally, 10-fold cross-validation is also used to verify that our results are unbiased.

The rest of this paper is organized as follows: the background and existing literature are explained in Section 2. The proposed approach, dataset, and other specifics are discussed in Section 3. Experimentation and assessment criteria are covered in Section 4. Section 5 consists of results and discussion. Finally, the conclusions and future work of this research are given in Section 6.

## 2. Background and Existing Literature

In the years ahead, SDN is likely to be the most promising networking model. An application plane, data plane, control plane, and respective APIs, i.e., southbound API and northbound API, make up SDN’s architecture. The communication between the applications and controller is based on the northbound interface. The functions of the southbound APIs include communicating with network virtualization protocols, switching fabric, and also a decentralized computing network. The SDN architecture separates the control plane from the application and data plane [8]. The control plane is a centralized and intelligent device that gives an overview of the underlying network. In addition, the control plane is a concentrated data processing and decision-making unit. It also can send data across the entire network. The data plane, on the other hand, represents the collection of SDN agents and the devices used for forwarding. Because the whole framework is dependent on the control plane, it is configurable and has the ability to expand its capabilities by incorporating further modules. As a result, SDN offers flexibility and creativity, and its detailed design is explained in [21]. All SDN controllers can extend different modules.

Because of this, the authors’ proposed detection technique is implemented on the control plane. The architecture and design of different SDN controllers are mostly the same; nevertheless, their functionality differs. The implementation language varies from controller to controller. Floodlight, for example, uses Java as its implementation language, while POX is written in Python. According to modern scientific evolution, the IoT has manifested competencies that touch almost every aspect of our life. Because of its ease of acquisition, IoT is vulnerable to a variety of security threats that must be handled. SDN is a powerful technology that offers a potential way out for IoT security and integrity.

In the past few years, scholars have shown a keen interest in DL and its applicability in a variety of fields, including vehicle production, law, and health care [22–24]. The DL techniques have improved the area of computer engineering through various applicabilities, which are practically employed in every industry, from medical appliances to self-driving cars. The deep neural network (DNN) models make use of the neural network architecture, which is why they are termed as deep neural networks [25–27]. These models are trained on a large amount of labeled data and to extract features from it without the need for human intervention. Additional DL applications include speech recognition software, fraudulent activity detection, image categorization, and intrusion detection. It can also be used to detect pedestrians, which reduces accidents. Different technological efforts have been made to address IoT’s vulnerable characteristics; nevertheless, SDN-based security solutions have shown to be the most effective [28]. Other cutting-edge technologies link with SDN to effectively fulfill the purpose under issue. The SDN blockchain integration is shown, which addresses all of the critical security apprehensions of IoT from an ultramodern standpoint. The primary ability of that amalgamation is the protection from DoS attacks, impersonating attacks, and routing attacks [29–32].

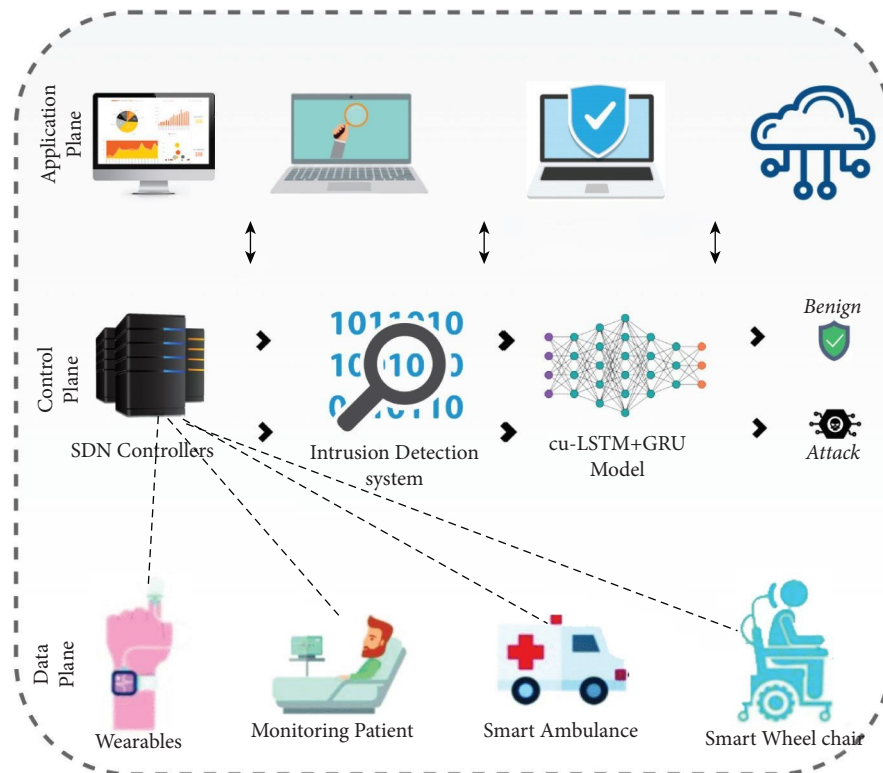


FIGURE 1: Proposed SDN-based model.

Furthermore, there is a lot of effort in the field of NIDS in SDN [33]. Another security model that should be discussed here is designed to protect the critical IoT ecosystem from many types of security attacks. The proposed scheme is a large-scale responsive atmosphere SDN-enabled blockchain-inspired solution. The model's performance is examined, and the positive results appear to make it an appropriate alternative for large-scale IoT networks [34]. SDN collaborates with the convolutional neural networks (CNN) to provide notable protection for IoT against a wide range of genuine issues. The tree of DDoS-based attacks is a warning indicator that communication in an IoT-based autonomous ecosystem may be disrupted. This behavior attracted the concentration of researchers, prompting the creation of an SDN-enabled CNN-based security architecture for IoT networks with limited resources. The proposed framework's most notable attribute is its ability to detect security threats quickly while using minimal network tools [7].

In terms of resource consumption, SDN-enabled security systems are thought to be outstanding. The SDN central controller's constitutional scheduling mechanism is always accompanied by exceptional network resource management. As a result, the attribute is passed down to SDN-enabled intrusion detection techniques, making it easier for IoT to satisfy defense frameworks while using the fewest resources possible [35]. In reference [36], the researchers presented a biometric mechanism to improve IoT security. The security of the system has been increased by an average of 96.82% using the suggested methodology. They used a combination of biometrics and coding. Based on experimental results, the

given solution enhances the security of the system by an average of 120.38%. By using biometric features and incorporating the findings of the evaluation, the risk of potential security issues occurring is reduced by 90.71%. Furthermore, because of IoT-specific service requirements (i.e., resource restrictions, low latency, flexibility, dissemination, and portability), attack detection differs dramatically from the previous approaches [36]. As a result, an adjustable, modular, dynamic, and cost-effective detection method against a variety of prevalent emerging cyber threats is critical for the IoMT networks. The authors of [37] used GRU-RNN for NIDS. They used the NSL-KDD dataset with six basic features and obtained an accuracy of 89%, which is insufficient for today's emerging security attacks.

In reference [38], an IoT-enabled healthcare system prototype-based framework is given. The solution makes use of a smart gateway design to make data storage and processing easier, and cloud-based analysis and decision-making. The security of this solution is determined by the operating system's security features and capabilities. The authors of [39] proposed a deep learning-based technique for detecting anomalies. CNN, LSTM, and MLP were employed in this system. Tshark and Wireshark were used to collect data for the experiment. In reference [40], the hierarchical architecture for usage in the domain of health is discussed, and the security of the data. Information relating to health data analysis is maintained separately in the cloud and fog infrastructure in this way. The MAPE-K-based model is also used in the solution to provide computations for executing various applications along with data encryption. In reference [41], the researchers suggested a DL

technique for flow-based intrusion based on a DNN. This framework used Snort (a network intrusion detection system) and Barnyard and obtained 85% accuracy. The authors of [42, 43] proposed a technique in SDN that relies on multilayer perception to overcome concerns with the botnet detection mechanism (MLP). Real data were used in the experiment, with a 98% accuracy rate. The authors proposed an RNN-based IDS in [44, 45] and used the NSL-KDD dataset for training. The analysis was carried out on the network traffic. For multiclass classification, this approach secured an accuracy rate of 81.29%. In reference [46], the authors described an intelligent SDN-based method for IoT intrusion detection. The researchers trained and experimented with deep learning classifiers on the CICIDS2017 dataset and improved detection accuracy.

### 3. Materials and Methods

This paper proposed an intelligent DL-driven threat detection technique for IoMT scenarios. This part covers our research approach, including the hybrid attack architecture, dataset description, proposed detection model, environmental setup, and metrics used for evaluation.

*3.1. Detection Technique and Network Model.* The SDN has grown in popularity as an embedded design during the last few years. The application plane of the SDN is designed to operate a wide range of apps and supply various services to end users. The control plane and the data plane are separated in the SDN design for simplicity and flexibility. On the other hand, the SDN's control plane is in charge of transmitting data, routing selections, and threat detection. Furthermore, the control plane improved the network's global view and main controller capabilities, making the collection of network data easier. To detect risks and exploitation in the IoMT environment, we propose Cu-LSTM+ GRU. The proposed model is placed in the SDN control plane, as shown in Figure 1. It is placed in the control plane for a variety of motives.

First and foremost, it is fully programmable and can also extend IoMT devices on the data plane. Second, SDN provides a solution for heterogeneity among IoMT devices and SDN controllers. Furthermore, the control plane can manage the primary IoMT devices in its data plane without depletion. The data plane is responsible for transporting data packets from the source to the destination and forwarding actual IP packets. The SDN framework and IoMT integration present a better solution to thoroughly monitor network traffic to detect intrusions, unauthorized events, and security attacks while being cost-effective and centrally controlled.

The Cu-LSTM+ GRU model is used in this strategy to detect advanced malware in the IoMT scenario. With better detection ratios and minimal false positives, the training and testing of the proposed model are performed by using the CICDDoS 2019 dataset. The proposed model consists of multiple layers, i.e., LSTM consists of 3 hidden layers with 600, 400, and 200 neurons while GRU consists of 2 layers of

300 and 150 neurons, respectively. For the activation function in the output layer, we employed softmax and ReLU in the other layers. The experimentation was carried out using 64 batch sizes until 20 epochs for better outcomes. The experiment is performed with the CUDA-enabled version. Furthermore, the proposed approach makes use of TensorFlow's backend and Python's Keras framework. A comparison is made with the proposed approach using the two classifiers. Cu-GRU+ DNN consists of 2 layers of GRU and 2 layers of DNN with 400, 300, 300, and 100 neurons. However, Cu-BLSTM has three layers with neurons of 400, 300, and 100, respectively.

*3.2. Dataset.* The selection of an adequate dataset is critical when evaluating the performance of threat detection schemes. The literature research reveals that different authors used different datasets for threat identification in such environments, such as NSL-KDD, KDD CUP99, and so on. Many of them lack the IoT support feature. Hence, the proposed work used an IoT-based dataset, i.e., CICDDoS2019 [47], which is publicly available. This dataset contains the most serious malware, such as DDoS and reflection attacks. Furthermore, the dataset is based on network flow and has IoMT supporting characteristics. The dataset contains more than 80 traffic features. The proposed model is concerned with 9 classes of the dataset. The details of the attacks and their instances are given in Table 1.

*3.3. Dataset's Preprocessing.* The following steps were used to preprocess the dataset in the proposed study. We initially identified all rows with NaN values and blank rows and further eliminated them completely, so the proposed model's performance and quality of data may not be affected. Using the label encoder, we next make the numeric values from all the non-numeric values, i.e., sklearn, because the DL algorithms mostly interpret numeric data. In addition, we used one-hot encoding on the output label to limit the odds of unexpected results, as model performance can be affected by category sorting. For data normalization, we used the MinMaxScaler, which improves the model's efficiency.

### 4. Environment/Experimental Setup

In our experiment, we used a graphic processing unit (GPU) and a Core i7-7700 processor for testing purposes. Furthermore, Python V3.9 and Keras have been used to train the suggested module. The experiment requirements, such as hardware and software requirements, are listed in Table 2.

*4.1. Metrics Used for Evaluation.* We assessed the suggested architecture's performance using standard assessment measures such as precision, recall, accuracy, and F1-score. In order to determine specific values (MCC), we have to calculate the true positive (TP), true negative (TN), false positive (FP), false negative (FN), false omission rate (FOR), and Matthew's correlation coefficient.

TABLE 1: CICDDoS 2019 details.

Attacks	Instances
Normal	56,600
DrDoS-MSSQL	2400
Dr-DoS	2350
DrDoS-SSDP	2368
PORTMAP	2496
UDP-lag	2300
SYN	2341
DrDoS-UDP	2600
WebDDoS	2365
Total	<b>75,820</b>

TABLE 2: Experimental setup.

Processor	I7 (3.33 GHz)
OS	Windows 10
RAM	16 GB
Language	Python
GPU	Geforce 1060
IDE	Spyder
Generation	8 <sup>th</sup>
Libraries	NumPy, TensorFlow, pandas, Keras, and scikit-learn

## 5. Results and Discussion

In this section, we have described the complete results of our proposed hybrid model (Cu-LSTM+ GRU). We also compared this model against two additional hybrid models, i.e., Cu-GRU+ DNN and Cu-BLSTM, and current methodologies in the literature, for a thorough performance review. The authors also performed a 10-fold cross-validation to show the unbiased results of the proposed model. The results are given in Table 3. Furthermore, the performance of our proposed model is assessed with the help of the standard metrics mentioned below.

**5.1. ROC Curve Analysis.** The effectiveness of an IDS can be evaluated using the critical metric known as ROC. True-positive (TPR) and true-negative (TNR) rates are associated, and the findings are plotted using ROC. The ROC curve for our approach is shown in Figure 2. The link between a true positive and a true negative is depicted in the following diagram. The figure depicts the efficacy of the proposed model.

**5.2. Confusion Matrix Analysis.** The classification model's output is shown in this evaluation matrix. The proposed model Cu-LSTM+ GRU accurately recognizes the classes based on the confusion matrix results. Figure 3 shows the confusion metrics for the proposed models proving that it successfully identifies the classes correctly and efficiently.

**5.3. Precision, Recall, Accuracy, and F1-Score.** The accuracy of a classifier demonstrates its efficiency and performance [48]. It indicates how many samples the suggested technique correctly identifies. The accuracy performance of the

proposed model is shown in Figure 4. This hybrid model has a 99.01% accuracy rate and a 98.80% recall rate. The records that are accurately identified reflect precision.

Furthermore, our suggested model has a precision of 99.04% and an F1-score of 99.12%, respectively. Complete detail of each fold is also given in Table 2 regarding the accuracy and other evaluation metrics. The per-class accuracy of all the three models is also provided in Table 4, proving the efficiency of the proposed model.

**5.4. FDR, FPR, FNR, and FOR Analysis.** We calculated the FDR, FOR, FPR, and FNR to adequately examine our proposed technique. Figure 5 shows the results. The FOR and FPR of Cu-LSTM-GRU have a value of 0.00172% and 0.00193%, whereas FNR and FDR are 0.00121% and 0.00164%, respectively. As a result, the proposed model, i.e., Cu-LSTM+ GRU, outperforms the other two models. Furthermore, Cu-GRU+ DNN shows better performance than Cu-BLSTM.

**5.5. MCC, TNR, and TPR Analysis.** To further assess the proposed model, we employed a confusion matrix to conduct an in-depth study of the MCC, TNR, and TPR analysis results. MCC, TNR, and TPR have values of 98.92%, 99.36%, and 99.13%, respectively. A closer examination of Figure 6 demonstrates that the proposed model outperforms the other two models.

**5.6. Speed Efficiency.** The testing time taken by our suggested method is demonstrated in Figure 7. We do not include the training phase because it was primarily performed offline. Testing is crucial when demonstrating the model's performance and efficiency. Our suggested hybrid

TABLE 3: 10-Fold results of Cu-LSTM+ GRU, Cu-GRU+ DNN, and Cu-BLSTM.

	Model	1	2	3	4	5	6	7	8	9	10
Accuracy (%)	Cu-LSTM+GRU	98.25	98.23	99.15	98.89	99.08	99.31	99.16	99.12	99.16	99.84
	Cu-GRU+DNN	97.56	97.21	97.86	97.54	98.54	98.57	99.15	98.81	98.62	98.86
	Cu-BLSTM	98.36	98.36	98.41	98.93	98.87	98.87	98.69	98.36	98.24	98.29
F1-score (%)	Cu-LSTM+GRU	98.24	98.63	99.68	99.06	99.06	99.25	99.19	99.34	99.08	99.68
	Cu-GRU+DNN	98.62	98.45	98.15	98.62	98.62	98.74	99.11	99.15	98.82	98.18
	Cu-BLSTM	98.94	98.91	98.29	98.29	98.68	98.15	98.19	98.81	99.16	99.43
Recall (%)	Cu-LSTM+GRU	98.96	98.92	99.26	98.61	98.21	98.61	98.89	98.97	98.69	98.92
	Cu-GRU+DNN	98.15	98.06	98.04	98.04	98.61	98.25	98.54	98.95	99.15	98.87
	Cu-BLSTM	98.15	98.16	98.16	98.85	98.71	98.06	98.15	98.64	98.64	98.86
Precision (%)	Cu-LSTM+GRU	98.16	98.68	99.14	99.14	99.32	99.36	99.86	99.51	98.91	98.34
	Cu-GRU+DNN	98.69	98.85	98.85	98.09	97.93	97.19	98.14	98.31	98.16	98.31
	Cu-BLSTM	98.19	98.96	98.48	98.48	98.86	98.46	98.69	99.05	99.17	98.78

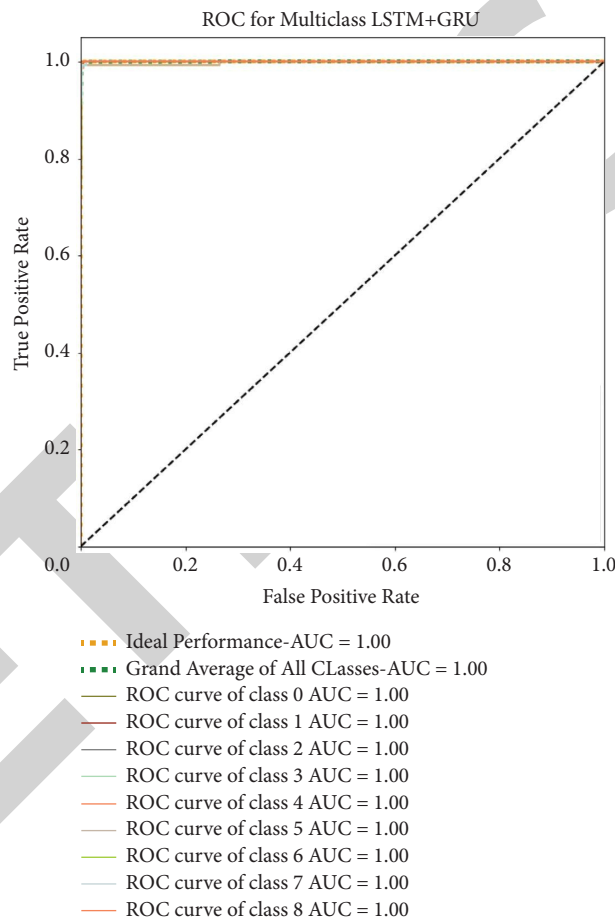


FIGURE 2: ROC curve of LSTM+ GRU.

techniques took only 19.35 ms to complete, which is a computationally efficient time. Cu-BLSTM, on the other hand, is computationally superior to Cu-GRU-DNN, having a testing time of 24.50 ms.

*5.7. The Comparison of Cu-LSTM+ GRU with the Existing Literature.* We compared the proposed method with the existing two hybrid DL models (Cu-GRU+ DNN and Cu-BLSTM) to demonstrate its efficacy. Both models were

evaluated using the same metrics and dataset, and the CICDDoS2019 dataset has been used to test and train all three models.

A comparison with other benchmark algorithms is also made. Table 5 shows a comparison of the suggested model to the current literature. The proposed model (Cu-LSTM+ GRU) clearly surpasses the existing literature regarding the accuracy, F1-score, precision, and speed efficiency. In addition, the suggested model's testing time is only 19.35 ms, which is much faster than previous benchmarks.

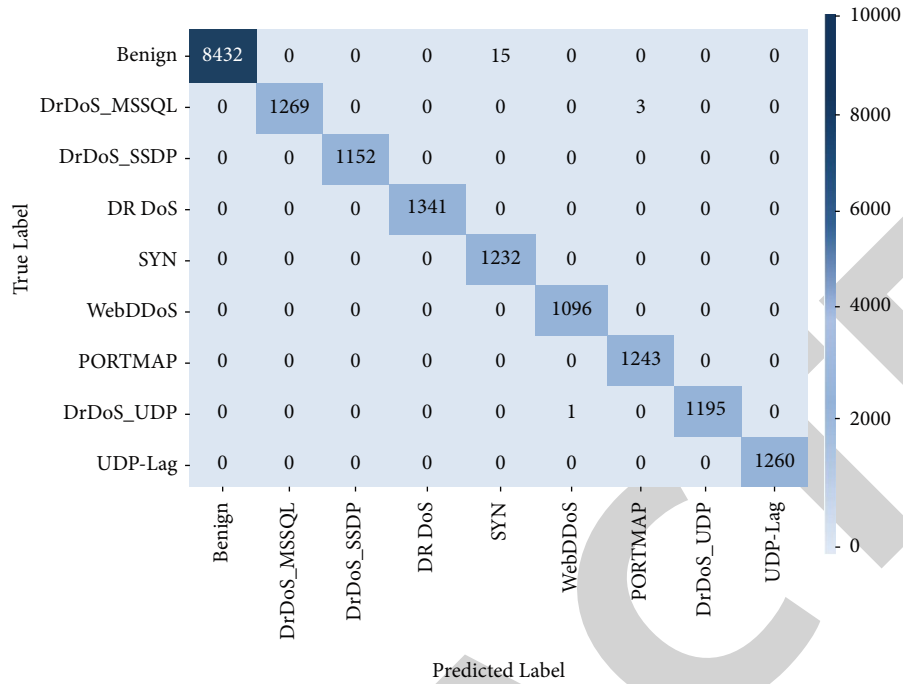


FIGURE 3: Confusion matrix of cu-LSTM+ GRU.

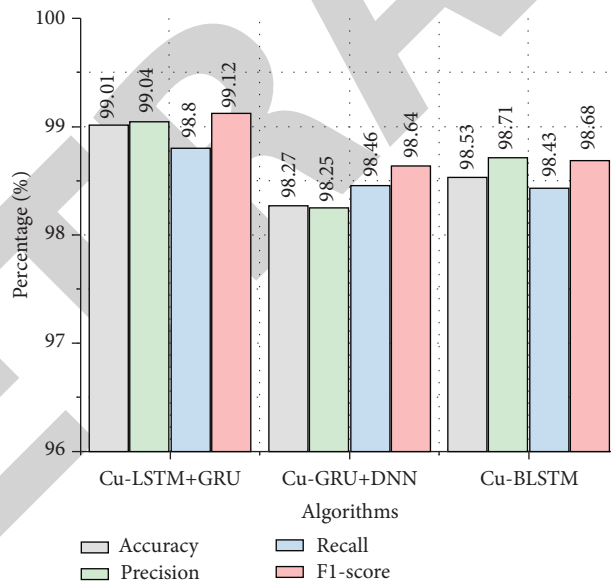


FIGURE 4: Overall comparison of the proposed model against Cu-GRU+ DNN and Cu-BLSTM.

TABLE 4: Per-class accuracy of the models.

Class	Cu-LSTM+GRU	Cu-GRU+DNN	Cu-BLSTM
Normal	99.84	98.86	98.93
DrDos-MSSQL	98.15	97.56	98.87
DrDoS-SSDP	99.12	98.54	98.29
DrDoS	98.23	97.86	98.36
SYN	98.25	97.21	98.24ss
WebDDoS	99.16	98.57	98.36
PORTMAP	99.31	99.15	98.93
DrDoS-UDP	99.08	98.62	98.41
UDP-lag	99.15	98.81	98.87



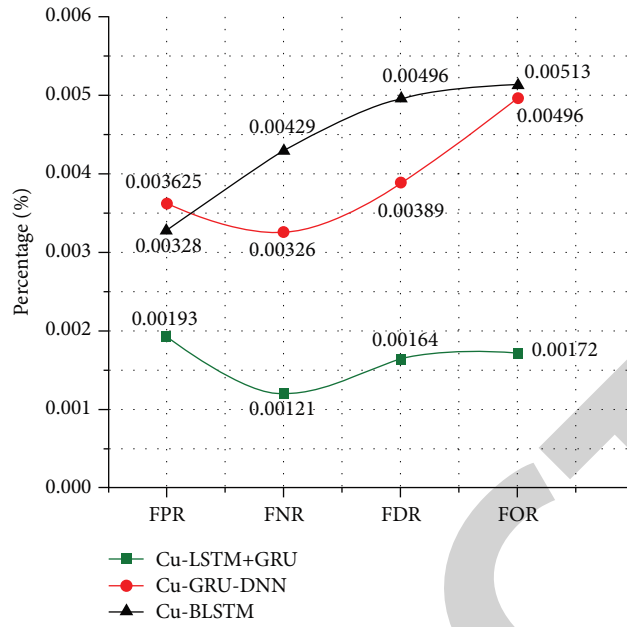


FIGURE 5: FPR, FNR, FDR, and FOR of the models.

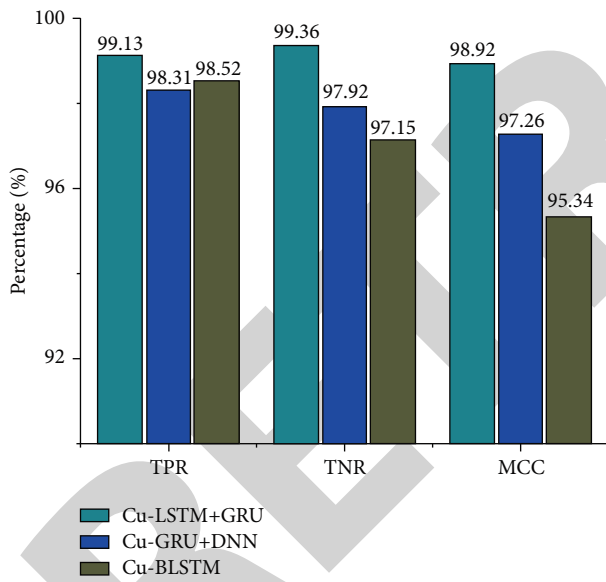


FIGURE 6: TPR, TNR, and MCC of the models.

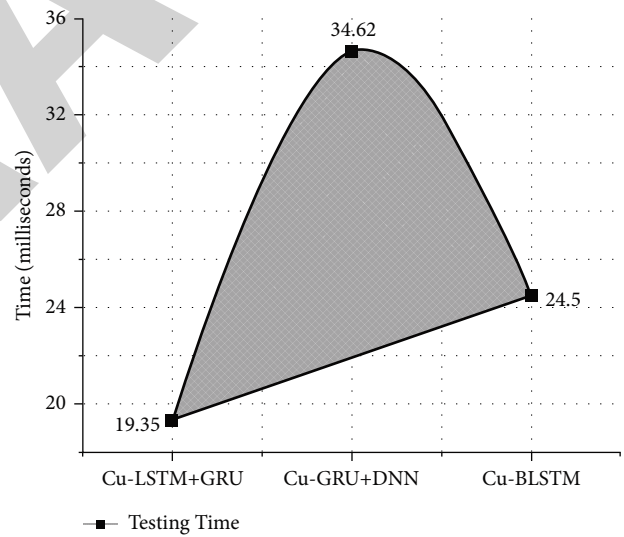


FIGURE 7: Testing time of Cu-LSTM + GRU, Cu-GRU + DNN, and Cu-BLSTM.

TABLE 5: Comparison with the existing literature.

Ref	Model	Accuracy (%)	Recall (%)	F1-score (%)	Precision (%)
Proposed	Cu-LSTM+ GRU	99.01	98.80	99.12	99.04
[49]	GRU-RNN	89.00	91.00	92.50	94.00
[17]	CNN	91.50	—	—	—
[50]	GRU-LSTM	87.90	77.90	80.60	83.50

## 6. Conclusions and Future Work

With the development of IoMT and E-health, the risk of cyber assaults has skyrocketed. These diverse devices make deploying traditional intrusion detection systems challenging in such environments. Therefore, the SDN paradigm provides a promising solution for protecting IoMT/E-health infrastructures. The proposed framework provides a quantitative, economical, and precise solution. A complete model test is run in combination with typical test metrics. We compared the result of the proposed model with two other classifiers that have been trained and evaluated under the same environment and with the current benchmarks. The proposed hybrid Cu-LSTM+ GRU model outperforms the current benchmark models with 99.01% accuracy and precision and F1-score of 99.12% and 99.04%, respectively. Furthermore, the computational complexity of the proposed model is very low, i.e., 19.35 ms. Despite its great performance, our proposed technique has a shortcoming that we intend to solve in the future, i.e., the proposed model would be more beneficial if it could identify insider threats.

In the future, we aim to use some other deep learning algorithms with blockchain to develop a new intrusion detection system for such environments. Finally, the authors endorse SDN-empowered, deep learning-based intrusion detection systems for the security of IoMT environments.

## Data Availability

Since the funding project is not closed and related patents have been evaluated, the simulation data used to support the findings of this study are currently under embargo while the research findings are commercialized. Requests for data, based on the approval of patents after project closure, will be considered by the corresponding author.

## Conflicts of Interest

The authors declare no conflicts of interest.





## References

- [1] J. Li, Z. Zhao, R. Li, and H. Zhang, "AI-Based two-stage intrusion detection for software defined iot networks," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 2093–2102, 2019.
- [2] G. Hatzivasilis, S. Othonas, I. Sotiris, and V. Christos, D. Giorgos and T. Christos, Review of security and privacy for the internet of medical things (IoMT)," in *Proceedings of the 2019 15th international conference on distributed computing in sensor systems (DCOSS)*, August 2019.
- [3] M. Asif, W. U. Khan, H. M. R. Afzal et al., "Reduced-complexity LDPC decoding for next-generation IoT networks," *Wireless Communications and Mobile Computing*, vol. 2021, pp. 1–10, Article ID 2029560, 2021.
- [4] A. B. Tufail, I. Ullah, W. U. Khan et al., "Diagnosis of diabetic retinopathy through retinal fundus images and 3D convolutional neural networks with limited number of samples," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 6013448, pp. 1–15, 2021.
- [5] R. Khan, Q. Yang, I. Ullah et al., "3D convolutional neural networks based automatic modulation classification in the presence of channel noise," *IET Communications*, vol. 16, no. 5, pp. 497–509, 2022.
- [6] M. Moradi, M. Moradkhani, and M. B. Tavakoli, "Security-level improvement of IoT-based systems using biometric features," *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 8051905, pp. 1–15, 2022.
- [7] I. Makhdoom, M. Abolhasan, J. Lipman, R. P. Liu, and W. Ni, "Anatomy of threats to the internet of things," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 2, pp. 1636–1675, 2019.
- [8] D. Javeed, T. Gao, M. T. Khan, and I. Ahmad, "A hybrid deep learning-driven SDN enabled mechanism for secure communication in internet of things (IoT)," *Sensors*, vol. 21, no. 14, p. 4884, 2021.
- [9] N. Ben-Asher and C. Gonzalez, "Effects of cyber security knowledge on attack detection," *Computers in Human Behavior*, vol. 48, pp. 51–61, 2015.
- [10] S. Garg, K. Kaur, N. Kumar, and J. J. P. C. Rodrigues, "Hybrid deep-learning-based anomaly detection scheme for suspicious flow detection in SDN: a social multimedia perspective," *IEEE Transactions on Multimedia*, vol. 21, no. 3, pp. 566–578, 2019.
- [11] W. Xia, W. Zhu, B. Liao, M. Chen, L. Cai, and L. Huang, "Novel architecture for long short-term memory used in question classification," *Neurocomputing*, vol. 299, pp. 20–31, 2018.
- [12] M. Al Razib, D. Javeed, M. T. Khan, R. Alkanhel, and M. S. A. Muthanna, "Cyber threats detection in smart environments using SDN-enabled DNN-LSTM hybrid framework," *IEEE Access*, vol. 10, pp. 53015–53026, 2022.
- [13] N. Koroniotis, N. Moustafa, E. Sitnikova, and B. Turnbull, "Towards the development of realistic botnet dataset in the Internet of Things for network forensic analytics: bot-IoT dataset," *Future Generation Computer Systems*, vol. 100, pp. 779–796, 2019.
- [14] B. K. Yousafzai, S. A. Khan, T. Rahman et al., "Student-performulator: student academic performance using hybrid deep neural network," *Sustainability*, vol. 13, no. 17, p. 9775, 2021.
- [15] A. B. Tufail, K. Ullah, R. A. Khan et al., "On improved 3D-CNN-based binary and multiclass classification of alzheimer's disease using neuroimaging modalities and data augmentation methods," *Journal of Healthcare Engineering*, vol. 2022, pp. 1–14, 2022.
- [16] S. Ahmad, T. Ullah, I. Ahmad et al., "A novel hybrid deep learning model for metastatic cancer detection," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 8141530, 14 pages, 2022.

- [17] J. Kim, J. Kim, H. Kim, M. Shim, and E. Choi, "CNN-based network intrusion detection against denial-of-service attacks," *Electronics*, vol. 9, no. 6, p. 916, 2020.
- [18] T. U. Khan, "Internet of Things (IOT) systems and its security challenges," *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)*, vol. 8, no. 12, 2019.
- [19] G. Acar, D. Y. Huang, F. Li, A. Narayanan, and N. Feamster, "Web-based attacks to discover and control local IoT devices," in *Proceedings of the 2018 Workshop on IoT Security and Privacy*, pp. 29–35, ACM, Budapest, Hungary, August 2018.
- [20] G. Aceto, V. Persico, and A. Pescapé, "A survey on information and communication technologies for industry 4.0: state of the art, taxonomies, perspectives, and challenges," *IEEE Commun. Surv. Tutor.* vol. 21, no. 4, pp. 3467–3501, 2019.
- [21] M. Du and K. Wang, "An SDN-enabled pseudo-honeypot strategy for distributed denial of service attacks in industrial Internet of Things," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 1, pp. 648–657, 2020.
- [22] I. Ullah, X. Su, X. Zhang, and D. Choi, "Simultaneous localization and mapping based on Kalman filter and extended Kalman filter," *Wireless Communications and Mobile Computing*, vol. 2020, Article ID 2138643, 12 pages, 2020.
- [23] A. Raza, H. Ayub, J. A. Khan et al., "A hybrid deep learning-based approach for brain tumor classification," *Electronics*, vol. 11, no. 7, p. 1146, 2022.
- [24] I. Ullah, Y. Shen, X. Su, C. Esposito, and C. Choi, "A localization based on unscented Kalman filter and particle filter localization algorithms," *IEEE Access*, vol. 8, pp. 2233–2246, 2020.
- [25] A. B. Tufail, I. Ullah, R. Khan et al., "Recognition of ziziphus lotus through aerial imaging and deep transfer learning approach," *Mobile Information Systems*, vol. 2021, Article ID 4310321, 10 pages, 2021.
- [26] I. Ahmad, I. Ullah, W. U. Khan et al., "Efficient algorithms for E-healthcare to solve multiobject fuse detection problem," *Journal of Healthcare Engineering*, vol. 2021, Article ID 9500304, pp. 1–16, 2021.
- [27] A. B. Tufail, Y. K. Ma, M. K. A. Kaabar et al., "Deep learning in cancer diagnosis and prognosis prediction: a minireview on challenges, recent trends, and future directions," *Computational and Mathematical Methods in Medicine*, vol. 2021, Article ID 9025470, pp. 1–28, 2021.
- [28] I. Alam, K. Sharif, F. Li et al., "A survey of network virtualization techniques for internet of things using SDN and NFV," *ACM Computing Surveys*, vol. 53, no. 2, pp. 1–40, 2021.
- [29] S. Ali, N. Javaid, D. Javeed, I. Ahmad, A. Ali, and U. M. Badamasi, "A blockchain-based secure data storage and trading model for wireless sensor networks," in *Proceedings of the International Conference on Advanced Information Networking and Applications*, pp. 499–511, Springer, Caserta, Italy, 2020, April.
- [30] I. Ullah, S. Qian, Z. Deng, and J. H. Lee, "Extended Kalman filter-based localization algorithm by edge computing in wireless sensor networks," *Digital Communications and Networks*, vol. 7, no. 2, pp. 187–195, 2021.
- [31] I. Ullah, X. Su, J. Zhu, X. Zhang, D. Choi, and Z. Hou, "Evaluation of localization by extended Kalman filter, unscented Kalman filter, and particle filter-based techniques," *Wireless Communications and Mobile Computing*, vol. 2020, Article ID 8898672, pp. 1–15, 2020.
- [32] X. Su, I. Ullah, X. Liu, and D. Choi, "A review of underwater localization techniques, algorithms, and challenges," *Journal of Sensors*, vol. 2020, Article ID 6403161, pp. 1–24, 2020.
- [33] D. Javeed, T. Gao, and M. T. Khan, "SDN-enabled hybrid DL-driven framework for the detection of emerging cyber threats in IoT," *Electronics*, vol. 10, no. 8, p. 918, 2021.
- [34] M. J. Islam, A. Rahman, S. Kabir et al., "Blockchain-SDN based energy-aware and distributed secure architecture for IoTs in smart cities," *IEEE Internet of Things Journal*, vol. 9, no. 5, pp. 3850–3864, 2021.
- [35] N. Mazhar, R. Salleh, M. Zeeshan, M. M. Hameed, and N. R.-I. D. P. S. Khan, "Real-time SDN based IDPS system for IoT security," in *Proceedings of the IEEE 18th International Conference on Smart Communities: Improving Quality of Life Using ICT, IoT and AI (HONET)*, pp. 71–76, IEEE, Karachi, Pakistan, October 2021.
- [36] R. Chaudhary, G. S. Aujla, S. Garg, N. Kumar, and J. J. P. C. Rodrigues, "SDN-Enabled multiattribute-based secure communication for smart grid in IIoT environment," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 6, pp. 2629–2640, 2018.
- [37] A. Molina Zarca, D. Garcia-Carrillo, J. Bernal Bernabe, J. Ortiz, R. Marin-Perez, and A. Skarmeta, "Enabling virtual AAA management in SDN-based IoT networks," *Sensors*, vol. 19, no. 2, p. 295, 2019.
- [38] A. M. R. Saharkhizan, "Exploiting smart e-Health gateways at the edge of healthcare Internet-of-things: a fog computing approach," *Future Generation Computer Systems*, vol. 78, pp. 641–658, 2018.
- [39] C. Li, Y. Wu, X. Yuan et al., "Detection and defense of DDoS attack-based on deep learning in OpenFlow-based SDN," *Int. J. Commun. Syst.*, vol. 31, no. 5, Article ID e3497, 2018.
- [40] I. Azimi, A. Anzanpour, A. M. Rahmani et al., "HiCH: hierarchical fog-assisted computing architecture for healthcare IoT," *ACM Transactions on Embedded Computing Systems*, vol. 16, no. 5s, pp. 1–20, 2017.
- [41] R. Vinayakumar, K. P. Soman, and P. Poornachandran, "Evaluation of recurrent neural network and its variants for intrusion detection system (IDS)," *International Journal of Information System Modeling and Design*, vol. 8, no. 3, pp. 43–63, 2017.
- [42] C. H. Huang, T. H. Lee, L. H. Chang, J. R. Lin, and G. Horng, *Adversarial Attacks on SDN-Based Deep Learning IDS System*, pp. 181–191, Springer, Singapore, 2019.
- [43] D. Javeed, U. M. Badamasi, T. Iqbal, A. Umar, and C. O. Ndubuisi, "Threat detection using machine/deep learning in IOT environments," *International Journal of Computer Networks and Communications Security*, vol. 8, no. 8, pp. 59–65, 2020.
- [44] F. Meng, Y. Fu, and F. Lou, "A network threat analysis method combined with kernel PCA and LSTM-RNN," in *Proceedings of the 2018 Tenth International Conference on Advanced Computational Intelligence (ICACI)*, pp. 508–513, IEEE, Xiamen, China, March 2018.
- [45] M. Latah and L. Toker, "Towards an efficient anomaly-based intrusion detection for software-defined networks," *IET Networks*, vol. 7, no. 6, pp. 453–459, 2018.
- [46] D. Javeed, T. Gao, M. T. Khan, and D. Shoukat, "A hybrid intelligent framework to combat sophisticated threats in secure industries," *Sensors*, vol. 22, no. 4, p. 1582, 2022.
- [47] I. Sharafaldin, A. H. Lashkari, S. Hakak, and A. A. Ghorbani, "Developing realistic distributed denial of service (DDoS) attack dataset and taxonomy," in *Proceedings of the 2019 International Carnahan Conference on Security Technology (ICCST)*, pp. 1–8, IEEE, Chennai, India, October 2019.

## Research Article

# An Intelligent Cost-Efficient System to Prevent the Improper Posture Hazards in Offices Using Machine Learning Algorithms

Jehangir Arshad <sup>1</sup>, Hafiza Mahnoor Asim,<sup>1</sup> Muhammad Adil Ashraf,<sup>1</sup>  
Mujtaba Hussain Jaffery <sup>1</sup>, Khurram Shabih Zaidi <sup>1</sup>, and Melkamu Deressa Amentie <sup>2</sup>

<sup>1</sup>Department of Electrical and Computer Engineering, COMSATS University Islamabad, Lahore 54000, Pakistan

<sup>2</sup>Department of Information Technology, Assosa University, Assosa 5220, Ethiopia

Correspondence should be addressed to Melkamu Deressa Amentie; [melkamu@asu.edu.et](mailto:melkamu@asu.edu.et)

Received 9 May 2022; Revised 27 June 2022; Accepted 14 July 2022; Published 18 August 2022

Academic Editor: Mohamed Abdelaziz

Copyright © 2022 Jehangir Arshad et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this research, an intelligent and cost-efficient system has been proposed to detect the improper sitting posture of a person working at a desk, mostly in offices, using machine learning classification techniques. The current era demands to avoid the harms of an improper posture as it, when prolonged, is very painful and can be fatal sometimes. This study also includes a comparison of two arrangements. Arrangement 01 includes six force-sensitive resistor (FSR) sensors alone, and it is less expensive. Arrangement 02 consists of two FSR sensors and one ultrasonic sensor embedded in the back seat of a chair. The K-nearest neighbor (KNN), Naive Bayes, logistic regression, and random forest algorithms are used to augment the gain and enhanced accuracy for posture detection. The improper postures recognized in this study are backward-leaning, forward-leaning, left-leaning, and right-leaning. The presented results validate the proposed system as the accuracy of 99.8% is achieved using a smaller number of sensors that make the proposed prototype cost-efficient with improved accuracy and lower execution time. The proposed model is of a dire need for employees working in offices or even at the residential level to make it convenient to work for hours without having severe effects of improper posture and prolonged sitting.

## 1. Introduction and Literature Review

The economic growth of any country around the globe is directly linked to the number of employees working in offices as it increases GDP, however, working nonstop sitting on office chairs can cause actual harm to mental and physical health. According to a survey, 60% of Americans experience real health problems because of the use of technology during the day, including insomnia, eye strain, headache, and more, out of which 30 percent have severe back pain and 27 percent have neck pain because of prolonged sitting in an improper posture [1]. Moreover, if a person sits on a chair in the wrong posture for a long time, it can cause more severe and painful diseases, such as pressure ulcers, back pain, spinal dysfunction, joint degeneration, rounded shoulders, and a potbelly. The wrong posture causes an increase in pressure, friction, and shear on the chair, and rubbing the chair with

the skin for a long time can tear the skin apart, resulting in severe pressure ulcers, leading to death.

In general, the office timings are eight hours, and people also work overtime or work longer than that without noticing their improper sitting patterns. For example, sitting eight hours on a chair with no physical activity can increase pressure at specific points in the body. The continuous rubbing of the skin on the chair surfaces results in pressure ulcers and other joint pains. Moreover, if the table height and distance from the table to a chair are not correct, the person has to bend toward his laptop, making a curve in his spine. These body curves can cause many hazards, such as back pain. Such improper postures, when taken for a prolonged period, can cause severe damage to health, and they must be prevented.

A good working environment is crucial in helping employees put their best foot forward. Sit-stand working

stations provide a good environment and reduce discomfort. Studies conclude that sit-stand working stations reduce perceived discomfort and increase productivity [2]. Sitting on the chair in an awkward posture or in one posture for a long time will cause harmful diseases, including pressure ulcers. The severe impacts of improper posture increased the importance of pressure recognition systems. There are many methods and techniques to detect the wrong posture to prevent getting improper posture and severe effects. One of the highest accuracy techniques to detect posture is FSR sensors and machine learning. Scientists and researchers have done much research on detecting posture accurately, mainly for wheelchair users. In addition to the traditional algorithms of the Naive Bayes classifier (NB), decision tree (DT), neural network (NN), multinomial logistic regression (MLR), and support vector machine, a system was developed to classify the sitting postures using different machine learning algorithms, such as the convolutional neural network (CNN) algorithm (SVM). A sensing cushion was created by installing a pressure sensor mat (8 8) within the children's chair seat cushion [3]. A pressure mapping device was used to measure the pressure in the right, one-sided prone, chin supported, and slumped sitting positions [4]. The WDI (XAP) found a statistically significant difference between the three incorrect sitting postures (one side prone, chin propped, and slumped sitting) and the correct sitting posture (one side prone, chin propped, and slumped sitting) [5]. An embedded device was utilized to acquire position-related information using a machine learning method termed dimensionality reduction (DR) and classification. The system includes a DR step based on the principal component analysis (PCA) [6]. For pose detection between the matrix stored (in the system) and new data obtained by pressure and distance sensors, K-nearest neighbours (KNN) classifiers are used [7–13]. Noninvasive optical fiber sensor architecture can be fitted to the bottom of a shoe for remote plantar pressure monitoring, which might be utilized in an IoT e-Health system to track people's health. The study looks at creating an optical fiber sensor multiplexed network (using fiber bragg gratings) to measure the distribution of foot plantar pressure during locomotion (walking) [8].

Picture data from the posture was captured using a film-type pressure sensor. The study enlisted the participation of twenty-six children, who were photographed in seven distinct poses. The authors employed a seven-layer convolutional neural networks (CNN) technique. In addition, the artificial neural networks (ANN) approach, one of the machine learning techniques, was utilized to compare classification accuracy [14–20]. The system (FPGA) is made of six flex sensors, an ADC board, and a machine learning algorithm based on a two-layer artificial neural network (ANN) developed on a Spartan-6 field programmable gate array [10]. The system achieves 97.78 percent accuracy with a floating-point evaluation and 97.43 percent accuracy with a 9-bit fixed-point implementation. The maximum propagation delay for the ANN and ADC control logic is 8.714 ns [11]. Ahmad et al. employed the J48 algorithm to identify the five types of sitting postures using the pressure readings of twelve pressure sensors as features,

achieving 99.47 percent experimental classification accuracy [12]. A center of pressure, contact area proportion, and pressure ratios are used in another article to identify five common trunk postures, two common left foot postures, and three common right foot postures. Lower-resolution mapping characteristics were compared to high-resolution sensor pressure mats on the backrest and seat pan features [21–26]. To recognize the postures of each body component, five distinct supervised machine-learning approaches are used [13]. Using a specific hardware system that interprets video in real-time using convolutional neural networks, a system based on the worker's postural detection is created, constructed, and tested. This device can identify the worker's neck, shoulders, and arms position and provide advice to help them avoid health problems caused by bad posture [12]. Many additional efforts, such as clinical implication assessment for diabetes mellitus [27], are built on machine learning techniques. There are also a slew of additional classification-based efforts to resolve health difficulties [28]. An intelligent way to forecast the accuracy of a model is to use an artificial neural network [29]. Multiple approaches are used in previous studies for the detection of posture, however, in this work, the cost is considered the main factor, and the higher accuracy of the system is achieved. Moreover, the proposed system is portable and multiple configurations of sensors are tested in this system [30]. Table 1 shows a comparison of prior research literature reviews. A system is designed to measure the variation in FSR sensor data on a chair according to the changing of the body sitting positions and the collection of that data.

- (i) Two arrangements have been used in this study to find the cost-efficient and higher accuracy arrangement of sensors according to the weight distribution of a sitting person on a chair
- (ii) Different machine learning techniques are used to clean the dataset to make it more efficient and useful
- (iii) Four machine learning algorithms are applied to detect the posture of a sitting person

The paper is divided into the following sections: designed system, materials and methods, hardware, software, configurations and specifications, data processing, results and discussion, conclusion, and future work.

## 2. Designed System

The sensors are connected to Arduino, and then the dataset is generated through serial communication, as shown in Figure 1. The data is stored in an excel file. In this way, data is stored, and then four different machine learning algorithms are applied to this data. This system supports mainly the table working in offices or at home. While working on laptops or computers in an office or at home, posture can easily be detected, and a warning of improper posture will be given on the laptop when that posture is taken for a prolonged period of time. The FSR sensors are very thin and small that they cannot be noticed, or no one can be annoyed

TABLE 1: Literature review.

Sr no	Paper	Type of sensors	No. of postures	Classifiers/software	Accuracy (%)
1	[3]	A film-type $8 \times 8$ FSR	5	(CNN), (DT), (SVM), (MLR), (NN), and (NB)	95.3
2	[11]	Inverse piezoresistive nanocomposite sensors	3	Three-layer BP neural network	98.75
3	[9]	Pressure sensor of film type	7	(CNN), (ANN)	97.5
4	[10]	An array of six flex sensors	7	Artificial neural network	97.8
5	[12]	16 sensors with 16 matrices	4	k-nearest neighbors (k-NN), support vector machines (SVM), random forest (RF), decision tree (DT) and LightGBM	99.03
6	[31]	Accelerometer, gyroscope, and magnetometer sensors	5	Naive Bayes, SVM, and KNN	99.90

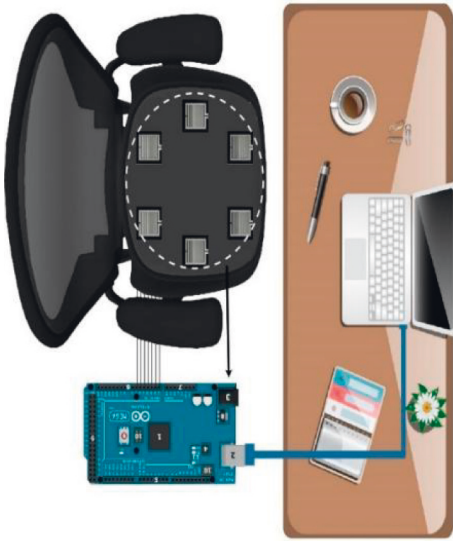


FIGURE 1: System diagram. FSR sensor data are being sent to the laptop through Arduino serial communication.

or disturbed by these sensors. Hence, it is an easy and efficient system to detect a person's posture when working on a table. It is a regular practice that people bend over to focus on the screen, and it can impact their health badly and have severe health hazards. Thus, this system makes posture detection more straightforward and efficient.

### 3. Materials and Methods

The process of posture detection can be done in various ways and methods, one of which is the approach discussed in this paper, which is the analysis and computation of pressure distribution data on a chair. In this approach, the force resistive sensors FSR are placed in a meaningful manner on a standard ordinary office chair, and when a person sits on it, the pressure distribution on the chair is measured by the sensors, and the posture is predicted using machine learning algorithms. In this paper, four machine learning algorithms are applied, including logistic regression, KNN, random forest, and Naïve Bayes, to predict the four improper sitting patterns, such as bending forward, bending backward, bending left, and bending right, as in Figure 2.

**3.1. System.** This study uses FSR (force-sensitive resistor) sensors that exhibit varying degrees of resistance to force directly proportional to the pressure exerted on them. Firstly, the FSR sensors are used to measure the pressure difference according to weight distribution on a chair. Then, this data is refined into a functional form to apply machine learning algorithms for the prediction or detection of the posture. The embedded system consists of two main parts, such as hardware and software.

**3.1.1. Hardware.** The weight distribution is evenly or unevenly divided on a chair when a person sits on it according to its sitting position. Hence, the pressure distribution on a chair is linked to a person's body mass and weight. Therefore, for the better capturing of pressure sensor data, the placement of FSR sensors must be meaningful. In this paper, two sensor placements are discussed to study the change in accuracy by changing the sensor arrangement. The two arrangements of sensors are as follows:

**(1) Arrangement 01.** In this arrangement, six FSR sensors are placed to cover the main pressure points of the human body, such as under the thighs, hip joints, and sides, as shown in Figure 3. When a person leans to the right, the right sensors have a higher-pressure value than the left ones and vice versa. In this way, the wrong posture is detected through the pressure distribution technique. The sensors work in a pair as the front two sensors are mainly used to detect the forward-leaning, and backward leaning is also detected by two sensors. When the pressure on the front two sensors increases, the person is sitting, leaning forward, and it is similar to backward leaning. Similarly, the left and right sensors are used to detect a person's left- and right-leaning positions. The chair used in this research is  $16 \times 16$  inches in size, and the vertical distance between the forward and backward pair of sensors is 5 inches, while the left-right sensors are embedded at edges, with a distance of 10 inches. The FSR sensor used in this study is approximately 1 inch long.

**(2) Arrangement 02.** The need for a second arrangement arises because of the higher cost of FSR sensors. In this approach, an ultrasonic sensor is used instead of FSR sensors for backward or forward-leaning detection.

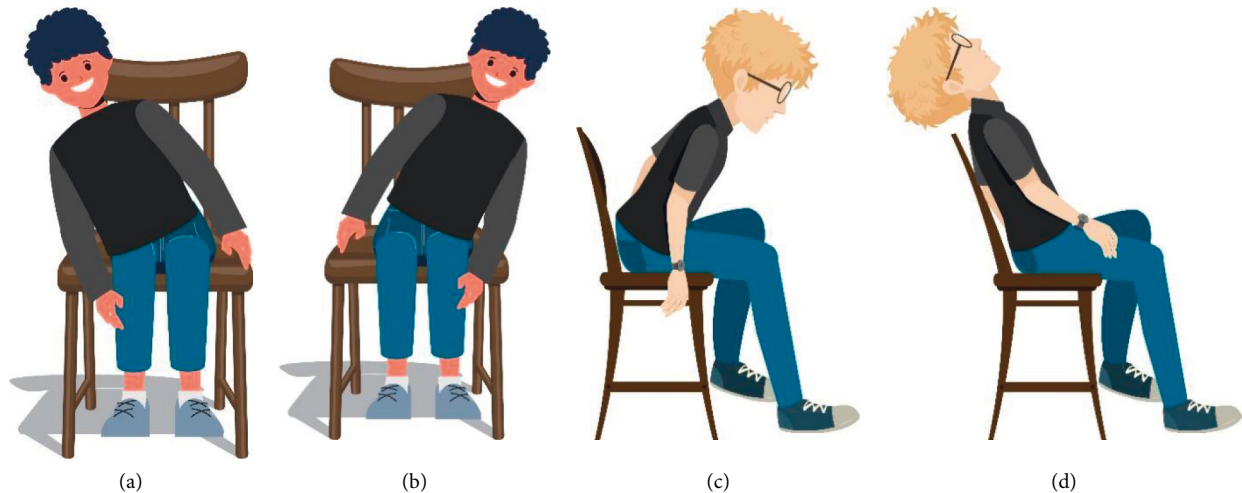


FIGURE 2: Four improper postures, namely left-leaning, right-leaning, forward-leaning, and backward leaning (from left to right).



FIGURE 3: Arrangement 01, six sensors arrangement on a chair.



FIGURE 4: Arrangement 02: two FSR and one ultrasonic sensor.

The optimal range for ideal posture is set on an ultrasonic sensor, and when the range exceeds, the forward-leaning posture is detected, and when the distance between the ultrasonic sensor and the person reduces to zero, it is predicted as a backward leaning position. Arduino mega is used for arrangement 01 because of the more significant number of FSR sensors, and Arduino UNO is used for arrangement 02 because of the smaller number of FSR inputs. For this reason, arrangement 02 is considered cost-efficient. Moreover, FSR sensors are reduced to more than half in arrangement 02. Figure 4 shows two FSR and one Ultrasonic sensor deployed in arrangement 2. A laptop processor is used for machine learning algorithms and Arduino serial communication. Table 2 shows hardware and specifications used for the proposed system.

The cost comparison of arrangement 01 and arrangement 02 is given in Table 3.

**3.1.2. Software.** The sensor data is sent to Arduino IDE via serial communication, and then Excel is used for storing the collected data. The Google collab notebook is used for

machine learning algorithms in this paper. The collab runs the code on local hardware. Thus, execution time may vary from system to system.

**3.2. Data Acquisition.** The data is collected at COMSATS UNIVERSITY Islamabad, Lahore campus, and students of different weights, heights, and body mass contributed to this study. The teachers and staff also contributed to dataset generation. Hence, the age ranges from 20 to 45 years. The number of samples collected is 635 different persons, and among them, 155 samples are forward-leaning, 145 are backward-leaning, 160 are right-leaning, and 175 are left-leaning, as shown in Figure 5(a). The BMI range of collected sample data varies from 16 to 35.

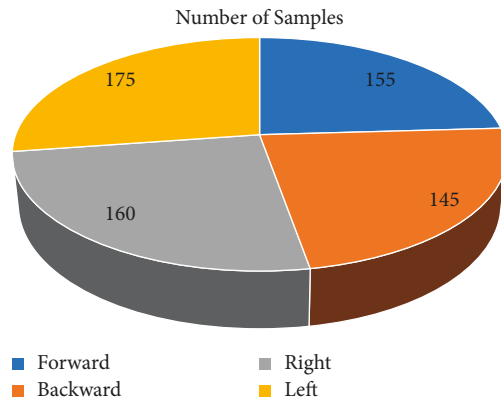
**3.3. Data Preprocessing.** Data preprocessing is an efficient technique to make raw data valuable and significant. Data preprocessing techniques refine the dataset into a functional form in this study. The repeated, corrupted, and null entries are eliminated from the dataset. Decreasing the irrelevant

TABLE 2: Hardware component details.

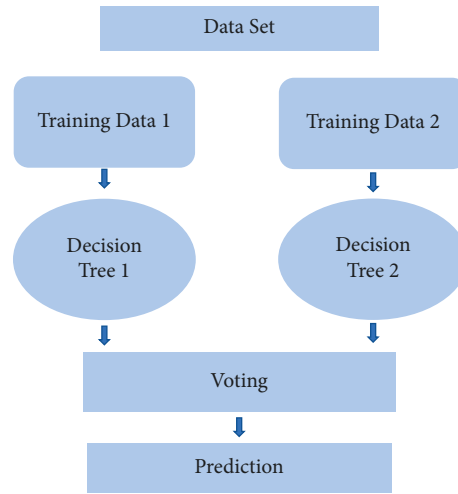
Hardware	Specification
Arduino mega	Arduino mega is a microcontroller kit for building digital devices based on the ATmega 2560. It has 54 digital input/output pins
Laptop	A laptop used for data collection and processing has the following specifications: intel core i5, 6 <sup>th</sup> generation, G3 2.40 GHz processor, and 8 GB ram
Force-sensitive resistance	A force-sensitive resistor (FSR) is a material that changes its resistance, when a force or pressure is applied, the output voltage varies from 0 to 5 V, depending on the amount of force applied to the sensor
Sonar sensor	Sonar sensors use ultrasonic sound waves to detect objects and distance. It sends an ultrasonic pulse out at 40 kHz, which travels through the air, and if there is an obstacle or object, it will bounce back to the sensor

TABLE 3: Cost comparison of arrangement 01 and arrangement 02 of sensors distribution.

Components	Arduino UNO	Arduino mega	FSR Sensors	Ultrasonic Sensors	Total
Arrangement 01	Quantity: 0	Quantity: 1	Quantity: 6	Quantity: 0	73.08\$
	Cost: 0	Cost: 35.88\$	Cost: 6.2 × 6 = 37.2\$	Cost: 0	
Arrangement 02	Quantity: 1	Quantity: 0	Quantity: 2	Quantity: 1	37.89\$
	Cost: 24.95\$	Cost: 0	Cost: 6.2 × 2 = 12.4\$	Cost: 0.54\$	



(a)



(b)

FIGURE 5: (a). Ratio of samples for postures, including forward, backward, left, and right. (b) Flow diagram 1: random forest working.

column is a significant step in removing the data and time columns from the dataset. Moreover, outliers that lead to the wrong prediction are removed. Another major step in data preprocessing is extracting null rows and columns and then removing them and placing them by their primary values. The smote analysis also removes the values where repetition exists to make the dataset more significant.

3.4. *Data Processing.* After the preprocessing of data, four machine learning algorithms are applied to detect the posture, and their accuracy and execution time are compared. Logistic regression predicts the output, which is present in categorical form. It can be either Yes or No, 0 or 1, true or False, etc., and it gives a value between 0 and 1. When applying logistic regression, the output is in categorical form. Logistic regression gives a value between 0 and 1, and it cannot go above this limit, forming an S-like curve. The

S-form curve is called the Sigmoid function or the logistic function [32]. The probability of a record belonging to the positive class given feature prediction by the logistic regression (9).

$$P = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_n X_n)}} \quad (1)$$

The other algorithm used in this study is KNN. Using KNN, the quick identification of the category or class is achieved [33]. The working of the KNN algorithm is as follows:

- (i) Selecting the number  $K$  of the neighbors
- (ii) Euclidean distance of  $K$  number of neighbors is calculated
- (iii) The  $K$  nearest neighbors from the calculated Euclidean distance



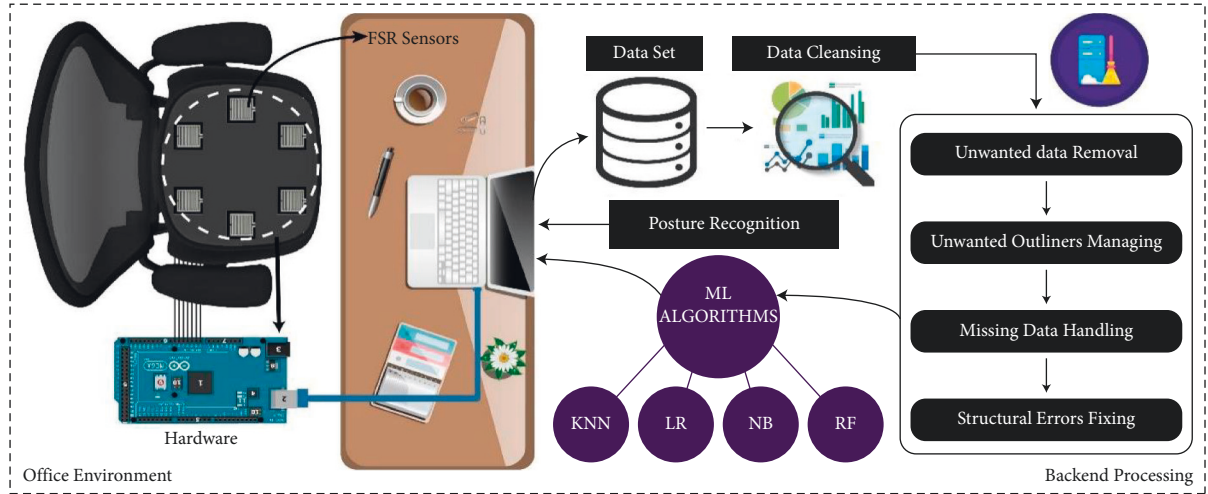


FIGURE 6: Complete system overview, including working and back-end environment.

- (iv) Counting the number of data points from  $k$  neighbors in each category
- (v) For which number of the neighbor is maximally assigning them a new data point

Distance functions used to calculate the distance from the nearest neighbor are Euclidean (2), Manhattan (3), and Minkowski (4).

$$\text{Euclidean} = \sqrt{\sum_{i=1}^k (x_i - y_i)^2}, \quad (2)$$

$$\text{Manhattan} = \sum_{i=1}^k |x_i - y_i|, \quad (3)$$

$$\text{Minkowski} = \left( \sum_{i=1}^k (|x_i - y_i|^q) \right)^{1/q}. \quad (4)$$

The working diagram of the random forest is as shown in a flow diagram in Figure 5(b).

Another algorithm used in this study is Naïve Bayes. In the naïve Bayes algorithm, the occurrence of a certain feature is independent of the occurrence of other features [34].

The working of the naïve Bayes algorithm is as follows:

- (1) Converting the given dataset into frequency tables
- (2) Likelihood table generation by finding the probabilities of given features
- (3) For calculating the posterior probability using the Bayes theorem

The highest accuracy provider algorithm is random forest. Both big and small data can be handled using random forest. The random forest is a combination of multiple decision tree algorithms [35]. Figure 6 illustrates a complete overview, including working and back-end environment of the proposed system.

KNN is a lazy learner that stores data at the time of classification, however, the advantage of using KNN is that it is straightforward to implement. The logistic regression classification algorithm gives about 86.78% precision and an accuracy of 83.34%, just in a training time of 0.029 s. Logistic regression produces accurate results using binary classification tasks, while the proposed system consists of multiclass classification. Random forest contains several decision trees consisting of various subsets in the given dataset. It takes an average to improve the predictive accuracy of that dataset. As the number of trees increases in the forest, it leads to higher accuracy and prevents the problem of overfitting. Random forest generates a predicting precision of 97.3%, an accuracy of 97.08%, and the training time for data is just 0.662 s. Naïve Bayes generates the predicting precision of 82.22%, the accuracy of 77.38%, and the training time for data is just 1.5 s. It is a probabilistic classifier, which means it predicts based on the probability of an object. The examples of the Naïve Bayes algorithm are spam filtration, sentimental analysis, and classifying articles. Arrangement 01 results and comparison of different algorithms is given in Table 4.

## 4. Results and Discussion

Machine learning algorithms are used in this research to predict four improper sitting postures. Two different arrangements are used to study the accuracy and cost-efficient system. This designed system has evaluated classification algorithms, including logistic regression, Naive Bayes, KNN, and random forest. Classifiers used in this system are judged based on their performance parameters, including precision, accuracy, and training time.

**4.1. Results of Arrangement 01.** For this arrangement, the KNN (k-nearest neighbor) algorithm predicting precision is 89.8%, accuracy is 93.97%, and execution time is about 0.552 s. KNN is a nonparametric algorithm. It never makes any assumption on underlying data. Figure 7 shows the plots of random forest algorithm for arrangement 01.

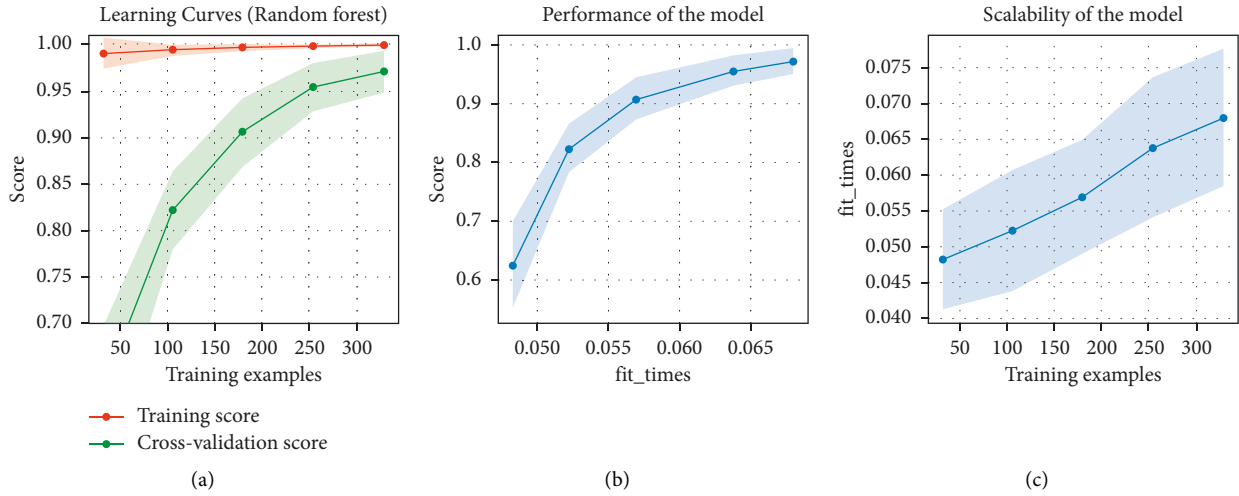


FIGURE 7: Random forest graphs: (a) learning curves, (b) performance of the mode, and (c) scalability of the model.

TABLE 4: Arrangement 01, machine learning algorithms comparison.

Algorithms	Precision (%)	Accuracy (%)	Training time (s)
Logistic regression	86.78	83.34	0.029
Naïve bayes	82.22	77.38	1.5
KNN	89.8	93.97	0.552
Random forest	97.3	97.08	0.662

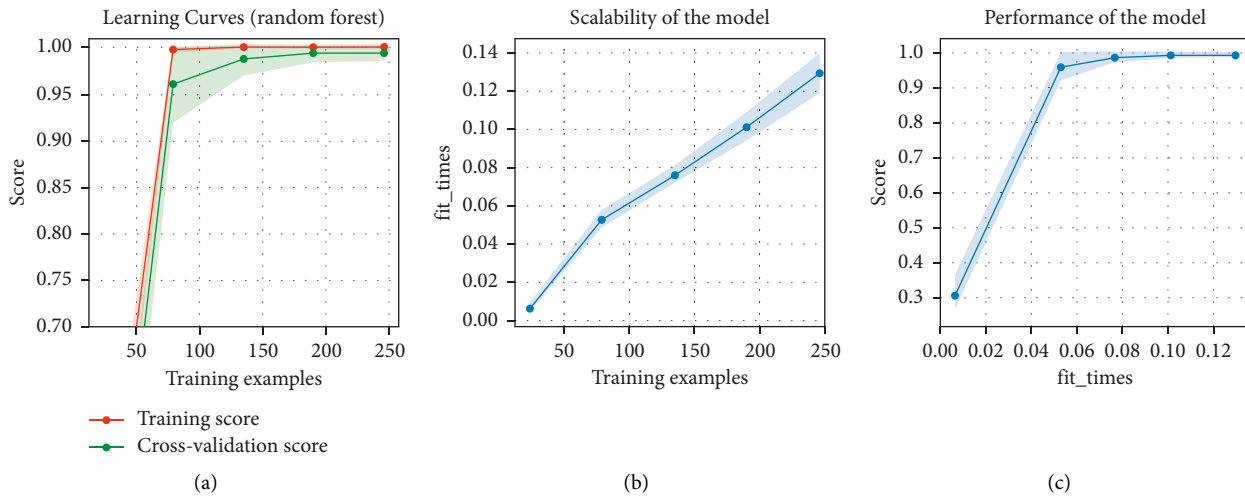


FIGURE 8: Table 3: arrangement 02: random forest graphs: (a) learning curves, (b) scalability of the model, and (c) performance of the model.

**4.2. Arrangement 02 Results.** In this arrangement of sensors, two FSR and only one ultrasonic sensor are used. For sonar configuration, KNN (k-nearest neighbor) algorithm predicting precision is 99.97%, accuracy is 99.97%, and its execution time is about 0.519 s. The Logistic Regression classification algorithm gave a precision of about 99.96% and an accuracy of 99.97%, just in a training time of 0.029 s. Random forest generates the predicting precision of 99.78%, the accuracy of 99.998%, and the training time for data is just 0.062 s. Naïve Bayes generates the predicting precision of

TABLE 5: Comparison of machine learning algorithms of arrangement 02.

Algorithms	Precision (%)	Accuracy (%)	Training time (s)
Logistic regression	99.96	99.98	0.029
Naive bayes	82.22	77.38	2.56
KNN	99.97	99.97	0.519
Random forest	99.78	99.998	0.062

82.22%, the accuracy of 77.38%, and the training time for data is just 1.5 s. Figure 8 and Table 2 show the plots of random forest algorithm for arrangement 02 and comparison, respectively. (Table 5)

## 5. Conclusion and Future Work

In this era of technology, everything and every solution is just a click away on technical gadgets and computers. Nevertheless, this technology stuff somewhere has a significant impact on the overall health factors of the human body. This study presents a solution to the improper sitting habits of any person who works a lot sitting at a desk. In this study, FSR sensors and an ultrasonic sensor are used to detect the sitting posture of a human body. Four machine learning algorithms are applied to the dataset generated on four different improper sitting postures: forward-leaning, backward-leaning, left-leaning, and right-leaning. The accuracy of 99.998% is achieved using the Random forest machine learning algorithm when applied to arrangement 02 data set: two FSR sensors and one ultrasonic sensor. The system can be used in offices and at home during desk work to prevent improper posture hazards and maintain fitness. This paper's vital scientific contributions are as follows: (1) compared to flex and textile sensor systems, the proposed technique reduces the number of sensors and processing complexity, resulting in lower hardware overhead. This energy-efficient and trustworthy continuous sitting position detection system will play a more critical role in reducing musculoskeletal disorders and protecting users' dignity. (2) The proposed solution does not have a drift problem, and hence, repetitive calibrations may be avoided. As a result, it is more dependable and power-efficient than the accelerometer-based system, resulting in a longer battery life. (3) The proposed approach provides a more private user experience than the camera-based system. For future work, the communication can be done wirelessly to send the sitting posture history to health professionals and analyze in a better/faster way to understand the medical history. Furthermore, the real-time posture detection on a smartphone can be done as it is easy to access and efficient to detect the posture while a person sits anywhere rather than specifically on an office chair or work desk.

## Data Availability

No data were used to support this study.

## Conflicts of Interest

The authors declare that there are no conflicts of interest.

## Authors' Contributions

J. Arshad, H. M. Asim, and M. A. Ashraf contributed to actualization, methodology, and initial drafting. M. H. Jaffery and K.S. Zaidi have done formal analysis software, and reviewed the manuscript. J. Arshad and M. D. Amentie contributed to visualization, validation, resources, and reviewing the draft. All authors read and approved the final version.

## References

- [1] M. M. Alam and E. Hamida, "Surveying wearable human assistive technology for life and safety critical applications: standards, challenges and opportunities," *Sensors*, vol. 14, no. 5, pp. 9153–9209, 2014.
- [2] T. Karakolis and J. P. Callaghan, "The impact of sit-stand office workstations on worker discomfort and productivity: a review," *Applied Ergonomics*, vol. 45, no. 3, pp. 799–806, 2014.
- [3] Y. M. Kim, Y. Son, W. Kim, B. Jin, and M. H. Yun, "Classification of children's sitting postures using machine learning algorithms," *Applied Sciences*, vol. 8, no. 8, p. 1280, 2018.
- [4] D. E. Lee, S. M. Seo, H. S. Woo, and S. Y. Won, "Analysis of body imbalance in various writing sitting postures using sitting pressure measurement," *Journal of Physical Therapy Science*, vol. 30, no. 2, pp. 343–346, 2018.
- [5] D. Van Eerd, C. Munhall, E. Irvin et al., "Effectiveness of workplace interventions in the prevention of upper extremity musculoskeletal disorders and symptoms: an update of the evidence," *Occupational and Environmental Medicine*, vol. 73, no. 1, pp. 62–70, 2016.
- [6] P. Rosero-Montalvo, D. Jaramillo, S. Flores, D. Peluffo, V. Alvear, and M. Lopez, "Human sit down position detection using data classification and dimensionality reduction," *Advances in Science, Technology and Engineering Systems Journal*, vol. 2, no. 3, pp. 749–754, 2017.
- [7] R. Gravina, P. Alinia, H. Ghasemzadeh, and G. Fortino, "Multi-sensor fusion in body sensor networks: state-of-the-art and research challenges," *Information Fusion*, vol. 35, pp. 68–80, 2017.
- [8] M. F. Domingues, N. Alberto, C. S. J. Leitão et al., "Insole optical fiber sensor architecture for remote gait analysis—an e-health solution," *IEEE Internet of Things Journal*, vol. 6, no. 1, pp. 207–214, 2019.
- [9] W. Kim, B. Jin, S. Choo, C. S. Nam, and M. H. Yun, "Designing of smart chair for monitoring of sitting posture using convolutional neural networks," *Data Technologies and Applications*, vol. 53, no. 2, pp. 142–155, 2019.
- [10] Q. Hu, X. Tang, and W. Tang, "A smart chair sitting posture recognition system using flex sensors and FPGA implemented artificial neural network," *IEEE Sensors Journal*, vol. 20, no. 14, pp. 8007–8016, 2020.
- [11] Z. Qian, A. E. Bowden, D. Zhang et al., "Inverse piezoresistive nanocomposite sensors for identifying human sitting posture," *Sensors*, vol. 18, no. 6, p. 1745, 2018.
- [12] J. Ahmad, J. Sidén, and H. Andersson, "A proposal of implementation of sitting Posture monitoring system for wheelchair utilizing machine learning methods," *Sensors*, vol. 21, no. 19, p. 6349, 2021.
- [13] V. K. Sinha, K. K. Patro, P. Pławiak, and A. J. Prakash, "Smartphone-based human sitting behaviors recognition using inertial sensor," *Sensors*, vol. 21, no. 19, p. 6652, 2021.
- [14] A. U. Rehman, R. A. Naqvi, A. Rehman, A. Paul, M. T. Sadiq, and D. Hussain, "A trustworthy IIoT aware mechanism as an enabler for citizen services in smart cities," *Electronics*, vol. 9, no. 6, p. 918, 2020.
- [15] A. Ur Rehman, A. Jiang, Abdul Rehman, and A. Paul, "Weighted Based Trustworthiness Ranking in Social Internet of Things by Using Soft Set Theory," in *Proceedings of the IEEE 5th International Conference on Computer and Communications (ICCC)*, pp. 1644–1648, Chengdu, China, December 2019.
- [16] Z. Hong, M. Hong, N. Wang, Y. Ma, X. Zhou, and W. Wang, "A wearable-based posture recognition system with AI-

- assisted approach for healthcare IoT,” *Future Generation Computer Systems*, vol. 127, pp. 286–296, 2022.
- [17] E. Hysenllari, J. Ottenbacher, and D. McLennan, “Validation of human activity recognition using a convolutional neural network on accelerometer and gyroscope data,” *German Journal of Exercise and Sport Research*, vol. 52, no. 2, pp. 248–252, 2022.
- [18] A. U. Rehman, R. Tariq, A. Rehman, and A. Paul, “Collapse of online social networks: structural evaluation, open challenges, and proposed solutions,” vol. 7, no. 11, pp. 1–6, Taipei, Taiwan, 2020.
- [19] M. H. Jaffery, M. A. Ashraf, A. Almogren et al., “FSR-based smart system for detection of wheelchair sitting postures using machine learning algorithms and Techniques,” *Journal of Sensors*, pp. 1–10, 2022.
- [20] A. Ur Rehman, A. Jiang, Abdul Rehman, A. Paul, S. Din, and M. Tariq Sadiq, “Identification and Role of Opinion Leaders in Information Diffusion for Online Discussion Network,” *Journal of Ambient Intelligence and Humanized Computing*, vol. 23, 2020.
- [21] T. Adane Gelaw and M. Tsighe Hagos, “Posture Prediction for Healthy Sitting Using a Smart Chair,” 2022, <https://arxiv.org/abs/2201.02615>.
- [22] Z. Fan, X. Hu, W. M. Chen, D. W. Zhang, and X. Ma, “A deep learning based 2-dimensional hip pressure signals analysis method for sitting posture recognition,” *Biomedical Signal Processing and Control*, vol. 73, Article ID 103432, 2022.
- [23] G. Farhani, Y. Zhou, P. Danielson, and A. L. Trejos, “Implementing machine learning algorithms to classify postures and forecast motions when using a dynamic chair,” *Sensors*, vol. 22, no. 1, p. 400, 2022.
- [24] K. Sornalakshmi, R. Venkataramanan, and R. Pradeepa, “Machine learning for human activity detection using wearable healthcare device,” in *Proceedings of the International Conference on Paradigms of Communication, Computing and Data Sciences*, pp. 711–724, Springer, Singapore, May 2022.
- [25] K. Bach, A. Kongsvold, H. Bårdstu et al., “A machine learning classifier for detection of physical activity types and postures during free-living,” *Journal for the Measurement of Physical Behaviour*, vol. 5, no. 1, pp. 24–31, 2022.
- [26] H. Katayama, T. Mizomoto, H. Rizk, and H. Yamaguchi, “You work we care: sitting posture assessment based on point cloud data,” in *Proceedings of the 2022 IEEE International Conference on Pervasive Computing And Communications Workshops And Other Affiliated Events (PerCom Workshops)*, pp. 121–123, IEEE, Pisa, Italy, 2022 March.
- [27] M. N. Sohail, R. Jiadong, M. M. Uba et al., “A hybrid Forecast Cost Benefit Classification of diabetes mellitus prevalence based on epidemiological study on Real-life patient’s data,” *Scientific Reports*, vol. 9, no. 1, 2019.
- [28] M. N. Sohail, R. Jiadong, M. U. Muhammad, S. T. Chauhdary, J. Arshad, and A. J. Verghese, “An accurate clinical implication assessment for diabetes mellitus prevalence based on a study from Nigeria,” *Processes*, vol. 7, no. 5, p. 289, 2019.
- [29] R. Rizwan, J. Arshad, A. Almogren et al., “Implementation of ANN-based embedded hybrid power filter using HIL-topology with real-time data visualization through node-RED,” *Energies*, vol. 14, no. 21, p. 7127, 2021.
- [30] J. Arshad, M. Aziz, A. A. Al-Huqail et al., “Implementation of a LoRaWAN based smart agriculture decision support system for optimum crop yield,” vol. 14, p. 827, 2022.
- [31] V. K. Sinha, K. K. Patro, P. Pławiak, and A. J. Prakash, “Smartphone-based human sitting behaviors recognition using inertial sensor,” *Sensors*, vol. 21, p. 6652, 2021.
- [32] C. Y. J. Peng, K. L. Lee, and G. M. Ingersoll, “An introduction to logistic regression analysis and reporting,” *The Journal of Educational Research*, vol. 96, no. 1, pp. 3–14, 2002.
- [33] G. Guo, H. Wang, D. Bell, and Y. Bi, “KNN Model-Based Approach in Classification,” *On The Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE*, vol. 2888, 2004.
- [34] Marlina, Leni, M. Muslim, A. P. Utama Siahaan, and P. Utama, “Data mining classification comparison (naïve bayes and C4. 5 algorithms),” *International Journal of Engineering Trends and Technology*, vol. 38, no. 7, pp. 380–383, 2016.
- [35] J. Ali, R. nullah Khan, N. Ahmad, and I. Maqsood, “Random forests and decision trees,” *International Journal of Computer Science Issues(IJCSI)*, vol. 9, 2012.

## Research Article

# A Cumulants-Based Human Brain Decoding

**Raheel Zafar** <sup>1</sup>, **Muhammad Javvad ur Rehman** <sup>1</sup>, **Sheraz Alam** <sup>1</sup>,  
**Muhammad Arslan Khan**,<sup>2</sup> **Asad Hussain** <sup>1</sup>, **Rana Fayyaz Ahmad** <sup>3</sup>, **Faruque Reza** <sup>4</sup>,  
and **Rifat Jahan** <sup>5</sup>

<sup>1</sup>Faculty of Engineering and Computer Science, National University of Modern Languages, Islamabad, Pakistan

<sup>2</sup>Department of Computer Science & Engineering, HITEC University, Museum Road, Taxila, Pakistan

<sup>3</sup>Centre for Intelligent Signal and Imaging Research, Universiti Teknologi PETRONAS, Seri Iskandar 32610, Perak, Malaysia

<sup>4</sup>Center for Neuroscience Services and Research, Universiti Sains Malaysia, Kubang Kerian 16150, Kota Bharu, Kelantan, Malaysia

<sup>5</sup>Department of Electronics and Telecommunication Engineering, Rajshahi University of Engineering & Technology (RUET), Rajshahi 6204, Bangladesh

Correspondence should be addressed to Rifat Jahan; 1604048@student.ruet.ac.bd

Received 25 April 2022; Revised 12 June 2022; Accepted 23 June 2022; Published 11 July 2022

Academic Editor: Ateeq Ur Rehman

Copyright © 2022 Raheel Zafar et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Human cognition is influenced by the way the nervous system processes information and is linked to this mechanical explanation of the human body's cognitive function. Accuracy is the key emphasis in neuroscience which may be enhanced by utilising new hardware, mathematical, statistical, and computational methodologies. Feature extraction and feature selection also play a crucial function in gaining improved accuracy since the proper characteristics can identify brain states efficiently. However, both feature extraction and selection procedures are dependent on mathematical and statistical techniques which implies that mathematical and statistical techniques have a direct or indirect influence on prediction accuracy. The forthcoming challenges of the brain-computer interface necessitate a thorough critical understanding of the complicated structure and uncertain behavior of the brain. It is impossible to upgrade hardware periodically, and thus, an option is necessary to collect maximum information from the brain against varied actions. The mathematical and statistical combination could be the ideal answer for neuroscientists which can be utilised for feature extraction, feature selection, and classification. That is why in this research a statistical technique is offered together with specialised feature extraction and selection methods to increase the accuracy. A score fusion function is changed utilising an enhanced cumulants-driven likelihood ratio test employing multivariate pattern analysis. Functional MRI data were acquired from 12 patients versus a visual test that comprises of pictures from five distinct categories. After cleaning the data, feature extraction and selection were done using mathematical approaches, and lastly, the best match of the projected class was established using the likelihood ratio test. To validate the suggested approach, it is compared with the current methods reported in recent research.

## 1. Introduction

The significance of neuroscience has become iconic in recent years due to the demand for intelligent systems in daily life. In the last decade, neuroscience becomes very famous due to its tremendous advances and applications. Neuroscientists from all over the world have progressed in this area; individual investigators also played a vital role in this field and worked in small groups on specific projects. Since neuroscience is an emerging field and has enormous applications,

and therefore, during the last decade, a huge amount of funding is provided. In 2013, US President approved a grant of USD 100 million to unlock the mysteries of the brain [1]. Similarly, in 2013, the European Union [2] also approved funding of 1 billion Euros for human brain projects. The main driving force in the field is the impact of brain disease on the population and the knowledge gaps in neuroscience, requiring a collaborative large-scale effort of scientists. The increasing demand for this field is due to its strength as it can be used in various fields. From medical to defense, it can be

used in any field including trials in court and as a basic research tool. Previously, it was used as evidence by the court in India [3], also used in various medical applications [4], and can be used for defense purposes [5, 6]. It can also be used to see the effect of the brain on various products. A blind test was done between Pepsi and Coke based on functional magnetic resonance imaging (fMRI) data to see which one is better [7]. The results showed that although Coke is a more valuable and successful brand, more people liked Pepsi. In Baylor College of Medicine, Houston performed another research in which fMRI data were used to see the brain activations for the taste of Pepsi and Coke [8]. The participants drank Pepsi or Coke and different parts of the brain light up depending on the cola being drunk. A linear regression analysis using behavioral preferences indicated the difference in brain responses evoked by Coke and Pepsi in the ventromedial prefrontal cortex. The average count was higher for labeled Coke compared to the labeled Pepsi. This shows that high activation was found when the participant drank their favorite cola. It is concluded that if acceptable accuracy can be achieved, then the prediction of the human brain can help in many applications including medical and health care.

In the above paragraph, the importance of prediction is discussed that how it can change our lives. Although several other applications are dependent on this field [9–11], the purpose of this paper is not to mention all the applications as they will be beyond the scope. The primary question that arises in the current situation is why we are not using this field for various applications as discussed above, for example, in court trials, and it is easy to predict from a suspect's brain measure whether he/she is involved in crime or not. The answer is it is not possible till now due to limited accuracy and reliability.

This research work comes under the field of computational neuroscience which uses theoretical neuroscience to validate and crack computational models. This should also be seen as a domain of theoretical neuroscience, but both areas are mostly related. The word mathematical neuroscience is used to emphasize the numerical character of the discipline occasionally [12, 13]. Computational neuroscience focuses on the description and physiology of biologically plausible neurons (and neural systems) as well as their dynamics, but biologically unreal models used for relativism, theory of control, cybernetics, quantitative psychology, machine learning, artificial neural networks, artificial intelligence, and computer education are not directly involved [14, 15]. While there is shared motivation and often no strict boundary within disciplines, paradigm abstractions depend on the complexity of the study and on the functionality by which biological structures are evaluated in computational neuroscience. In the current study, a method is proposed that can improve the accuracy and hence make this field more reliable. There are different steps from experiment design to the prediction of the class including data collection, feature extraction, and selection as shown in Figure 1. Different studies focused on different steps, but the common goal was to achieve better accuracy [16, 17]. For example, a better experiment design can improve prediction accuracy, and

similarly, by collecting quality data, accuracy can be improved. In short, the current research is mainly focusing on the analysis of fMRI data instead of depending on new hardware so in this study we have also focused to improve the accuracy by introducing a novel method. Since fMRI is a mature modality in neuroscience and is used for two decades, a lot of literature is available, and various techniques are used for better results. The primary focus of these studies was on experiment design and specific brain regions instead of new statistical techniques like many studies that used a support vector machine (SVM) for classification and took features from the region of interest (ROI) [17–20]. Therefore, the target of this study is to introduce a novel method in neuroscience and compare the results with the existing methods.

Prediction of the human brain is always a challenging task for neuroscientists due to its complex structure and rapid changing behavior. In the initial years of research, univariate-based approaches were very common and provided acceptable results [21, 22]. Univariate is the simplest form of analyzing data in which the data being analyzed contains only one variable. As univariate has a single variable, it does not deal with relationships or causes. Over time, improvements were done not only in hardware but new methods were also introduced for the analysis of complex brain data. Instead of univariate analysis, brain mapping and multivariate pattern analysis (MVPA) were introduced which improved the previous results significantly [17, 23, 24]. In MVPA, the primary focus is on distributed patterns of activity for analysis and comparison. In this way, the differences between different brain conditions can be detected with higher sensitivity.

In the initial years, functional magnetic resonance imaging (fMRI) was the most common, reliable, and widely used approach for the collection of brain data [22, 25], and in most of the initial studies, the data were extracted only from a particular brain region like occipital region for any visual task or from a temporal region in case of a cognitive task. As time passes, other modalities were also extensively used in neuroscience which provided good results such as an electroencephalogram (EEG), magnetoencephalography (MEG), and functional near-infrared spectroscopy (fNIRS) [26, 27]. In fMRI, brain activity is measured by identifying the changes that occurred due to the flow of blood. The fMRI uses the blood oxygen level-dependent (BOLD) contrast which was initially discovered by Seiji Ogawa [28]. In case of any neuronal activity, more blood is needed in that region which later flows and increases the blood level in that region. Due to blood flow, more BOLD contrast can be observed in that region. Since, in fMRI, brain activity is measured using blood flow, it is an indirect method to measure brain activity. Functional MRI is an indirect method of measuring brain activity patterns, but still the best among other modalities due to its higher spatial resolution. On the other hand, in EEG, the direct neuronal activity can be measured as its time resolution is in milliseconds, which is good compared to fMRI as it has in seconds.

The collection of brain data is itself a difficult task including experiment design and conduction. However, after the recording of brain data, the primary goal is to clean the data and find a significant difference in the temporal and

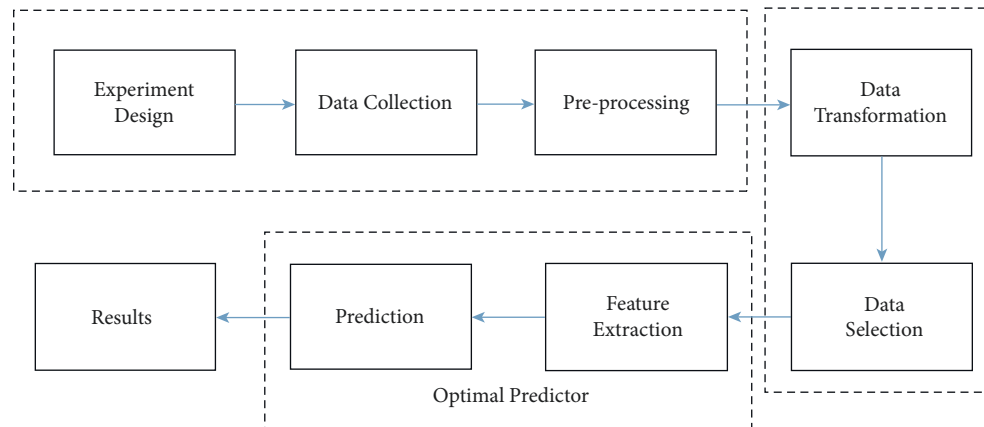


FIGURE 1: System model for the analysis of fMRI data.

spatial characteristics of the data, which contrasts the various states as precisely as possible. The goal of neuro-computing is to achieve this precise contrast, so a variety of simple and traditional statistical methods such as grand averages and region of interest were used in building the good model [22, 29, 30]. Currently, these methods are not common due to limited accuracy and reliability in comparison with new and complex statistical methods [20, 31, 32]. Although it is not possible to measure and analyze the billions of neurons present in the brain and none of the current modalities can do that, improvement is always demanding and challenging. Moreover, it is not easy to update the hardware regularly. That is why presently the focus of neuroscientists is on computations based on statistical and machine learning techniques which are applied in various fields, to improve the results instead of waiting for new technology [33–36].

Various research groups are working on neuro-computing, and improvements have already been observed both in terms of hardware and computational techniques, but this field still has a lot of potentials. Neuroscientists also used both fMRI and EEG together (simultaneously) in some studies [37–39] for improved results. In short, this field still has a chance of improvement in many ways. In hardware improvements, 7 Tesla fMRI machines and 512 channel EEG caps are available, and researchers are focusing on statistical and mathematical approaches to bring advancement in existing techniques with the help of advanced hardware. Many novel approaches are observed during the last decade, especially for feature extraction, feature selection, and prediction [26, 40]. Various projection and dimension reduction methods also emerged for the extraction of significant data. The common algorithms used for feature extraction are principal component analysis (PCA), independent component analysis (ICA),  $t$ -test, and many others [17, 41–43].

Apart from the above-discussed techniques, machine learning plays a vital role in the prediction of brain activities as the brain consists of a huge amount of data. Different machine learning and prediction techniques are used in this regard such as linear discriminant analysis (LDA), support vector machines (SVM), Naïve Bayes, Bayesian, and many others [42, 44–51]. Each of them took part in better results;

however, SVM is the most common in neuroscience compared to others. The likelihood ratio test (LRT) is also a statistical test used for comparing the best fit of two statistical models. LRT has mostly been used in likelihood ratio-based score fusion (LRBSF) which is a mature and widely used technique, especially in biometric systems [52], and however, it is rarely used in the neuroscience research area. In an existing study, [53], it was shown that LRBSF can produce better results in terms of accuracy for both fMRI and EEG data sets as compared to many existing methods, especially SVM.

*1.1. Contribution of the Research.* In this manuscript, a hybrid method is proposed. In the proposed algorithm, time series mathematical equations of various cumulants are used along with the likelihood ratio-based score fusion (LRBSF) method. The results are compared with the recent state-of-the-art methods. In the proposed method, the initial  $\beta$  values are taken as features that are found using GLM. These features are further refined with various orders (up to fourth) of cumulants and finally used likelihood ratio-based score fusion (LRBSF) which helped in better results with less time. In LRBSF, the number of features plays a great role, and limited features increase the performance by reducing the processing time of the existing method other than the prediction results.

*1.2. Organization of the Article.* The rest of the paper is organized as follows. Section 2 presents the experimental methodology of this paper. Section 3 deals with the statistical techniques which are mainly used in this research article. Section 4 represents the results based on the experimental setup, and a detailed discussion is given in the section. In the last Section 5, a conclusion of this research work is presented.

## 2. Experimental Methodology

The motivation behind the presented research work is to propose a method based on cumulants that can help in improving the accuracy using the fMRI data because the

implementation of neuroscience applications is dependent on higher accuracy. The reliability of any method is dependent on accuracy; if the accuracy is higher, the method is considered accurate and reliable. Cumulants are widely adopted in classification problems such as modulation recognition [54, 55]. To the best of our knowledge, this approach is seldom explored for fMRI analysis. Therefore, a novel prediction approach of cumulants-driven likelihood ratio-based score fusion (LRBSF) along with multivariate pattern analysis (MVPA) is presented in this research article to optimize the prediction results based on the fMRI data set.

For this study, the experiment was designed in e-prime [56], and fMRI was collected using a 3 T Philips machine and preprocessed using standard preprocessing methods in SPM [57]. During analysis, raw data were taken as features that were followed by feature selection done using a *t*-test. Multivariate pattern analysis (MVPA) has been used with fMRI data to extract information from distributed activation patterns of the brain. The significant features were still in large numbers, so different orders of cumulants were applied to reduce the number of features. Finally, the likelihood ratio test (LRT) was applied to those selected features and found the prediction accuracy.

The main steps of fMRI data analysis include experiment design, data collection, preprocessing, data transformation, data selection, feature extraction, prediction, and finally, results which are shown in Figure 1, and details are given in the following section.

*2.1. Experiment Design.* The initial step is the experiment design which helps in the extraction of data from the brain. Better experiment design can collect valuable and numerous data from the brain like event-related design, which helps in measuring the rapid change of brain-behavior. For the fMRI study, experiment design is an important factor as fMRI scans are expensive and time-consuming and need human resources. That is why experiment design should be carefully designed which can extract sufficient data for further analysis. The best experiment design can maximize the contrast of interest.

*2.2. Experimental Data Collection.* After the experiment design, data collection is a key step in which extensive care is required to retrieve useful data. The main goal during data collection is to record the data according to the task with the minimum artifacts or noise. In fMRI, data collection requires special consideration. The collected data should have strong physiological signals so that the discrimination between the brain states can be done easily along with the safety measures due to human involvement. For the experimental setup, healthy subjects were selected, and all were Hospital Universiti Sains Malaysia (HUSM) students. Their ages were between 18 and 26. The approval of the study protocol was taken from the Human Research Ethics Committee (HREC), Universiti Sains Malaysia (USM) under IRB Reg. No: 00004494, FWA Reg. No: 00007718, and USM issuing letter number is USMKK/PPP/JEPeM[257.3.(4)]. Therefore, it does not require registration with the

International Committee of Medical Journal Editors (ICMJE). The consent from all participants (subjects) was obtained before the start of the experiment. For safety purposes, the data were recorded in the company of the doctor at HUSM. A brief training session was arranged for each participant before the data collection.

Subject selection and machine parameters should precisely be chosen. Moreover, the duration of the session should not be too long. In this study, the data were collected from 12 subjects, and 8 were used for the analysis as data of 4 subjects were excluded due to artifacts and low accuracy. A total of 260 images were shown to the subjects in three different sessions, and all the images were grayscale images. These images were taken from the study [53] and are available online. These images were belonging to five different categories as described earlier means a picture of a human or animal or fruit or natural scene or building was presented to the subject. The subject should see the image carefully and recognize the category or class of the image. The target of the study is to see and find the differences in brain activities for various types of images significantly, in other words, to know whether the brain behaves differently for different categories.

Functional MRI data were collected using 3 Tesla machines with response time (TR) = 2000 ms, echo time (TE) = 30 ms. After every 2 second, 35 slices of the brain were taken. The anatomical data of 5 mins were separately taken which was later used during the registration of functional images with a structural MRI image. The preprocessing of the data was done using SPM 8. All the gradient echo-planar imaging (EPI) images were realigned to the first image as it has minimum head movement. The normalization was done using the Montreal Neurological Institute (MNI) space, and spatial smoothing was done with a voxel size of  $3 \times 3 \times 3$  mm. The final dimension of each fMRI image is  $63 \times 53 \times 46$ . The proposed method is trained and tested along with other methods that are used as the baseline compared to the proposed method. The reported results are with 8 subjects, and the average accuracy of them is mentioned in the result section.

*2.3. Preprocessing.* Preprocessing also plays a vital role during the analysis as it detects and reduces most of the artifacts and noise from the data. The main steps involved in preprocessing are slice timing correction, realignment, coregistration, segmentation, spatial normalization, and smoothing. In slice time correction, the differences in slice acquisition times are corrected. This correction is required to make the data on each slice correspond to the same point in time. The realignment process detects and corrects the motion of the subject inside the scanner during recording. During the process of coregistration, the functional images (T2) are aligned with the anatomical MRI images also known as T1 or structural images. In segmentation, the gray matter, white matter, and cerebral spinal fluid are separated and can be seen separately. Since every brain size is different, normalization is done to put the different scanned images into a standard Montreal Neurological Institute (MNI) template.



The standard MNI template is based on the average of MRI scans of healthy subjects. In smoothing, average values of the neighbor voxels are found for minimal noise. This is the last step of preprocessing which blurs the fMRI image by giving different weights based on the Gaussian kernel.

**2.4. Data Transformation.** After preprocessing, data transformation is done which is the initial stage of analysis. Data can be transformed in various ways, depending on the analysis procedure. In data transformation, the data are converted from one format to another format which may be more suitable for analysis. The new format can be a new destination system. The fMRI data are a 4-D data set that can be transformed into a 1-D data vector or can analyze as 2-D data with a different number of slices against the time.

**2.5. Data Selection.** The next step is data selection which has a convincing role in analysis, as it helps in extracting most of the significant information from the brain. The brain's slight behavior is changed during the task so data selection is necessary as most of the fMRI data slices do not have any contrast information. The size of fMRI data is large and less significant, so data selection becomes important in fMRI data. In the literature, various algorithms and methods are used for data selection like  $t$ -test, analysis of variance (ANOVA), entropy, Bhattacharyya distance, and others. In this study,  $\beta$  values found using  $t$  stats are used as initial features which are later refined using cumulants.

**2.6. Feature Extraction.** Feature extraction is an important part in which significant features should be extracted to make the analysis more valuable and reliable with improved accuracy. In fMRI data analysis, various features are already used such as region of interest, temporal-based feature selection, Dynamic causal modeling, features found using convolutional neural network, and many others. This essential part mostly uses statistical and mathematical methods to correlate different brain areas and helps in finding differences between the tasks. In the current study, various order cumulants are used as features that are given to the classifier for prediction.

**2.7. Prediction.** In this final step, prediction algorithms are used which help in determining the right classes for the testing data. Various algorithms and machine learning methods are used for this specific purpose such as support vector machine (SVM), linear discriminant analysis (LDA), Softmax, and others. The outcome of optimal predictor is in the form of accuracy based on true positives and true negatives which is later compared with various existing methods for validation.

### 3. Statistical Methodology

The proposed methodology of the paper is described as follows.

**3.1. General Linear Model.** The  $t$ -stats were found using a  $t$ -test from the raw data used as features for further analysis. These  $t$ -stats are found using the general linear model (GLM). The general linear model (GLM) is similar to statistical analyses [58], but it is also completely ideal both for various contextual and multifaceted variables. GLM is suitable for carrying out all parametric statistical tests with a dependent variable, like some factory configuration ANOVA, and designs with a combination of multiple variables (covariance analysis, ANCOVA). GLM is a key instrument for fMRI data analysis since being introduced by Friston and colleagues into the neuroimaging community [59] thanks to its versatility to integrate numerous, quantitative, and qualitative independent variables. In terms of a linear combination of the explanatory variables and an error term, a general linear model describes the response variable  $Y_j$ .

$$Y_j = x_{j1}\beta_1 + x_{j2}\beta_2 + \dots + x_{jL}\beta_L + \epsilon_j. \quad (1)$$

The  $\beta$  parameters here are undisclosed to be estimated.  $x_{jL}$  is the explanatory variable, and  $\epsilon$  is the error term that is arbitrary, with zero mean and variance  $\sigma$ . They are distributed independently and equally normally (i i d). The  $x_j$  must be calculated  $j = 1, 2, \dots, j$  and  $L = 1, 2, \dots, L$  for each observation. In a single session, there are 400-time points mean value of  $J=400$ , and 9 or 12 columns mean value of  $L=9$  or 12, according to the experiment performed. GLM can be written in matrix form which helps in deriving least squares parameter estimation. Equation (1) can be expanded as follows:

$$\begin{aligned} y_1 &= x_{11}\beta_1 + x_{12}\beta_2 + \dots + x_{1L}\beta_L + \epsilon_1, \\ &:= := :=, \\ y_j &= x_{j1}\beta_1 + x_{j2}\beta_2 + \dots + x_{jL}\beta_L + \epsilon_j, \\ &:= := :=, \\ y_J &= x_{J1}\beta_1 + x_{J2}\beta_2 + \dots + x_{JL}\beta_L + \epsilon_J, \end{aligned} \quad (2)$$

where  $Y$  is the observing column vector  $Y = y_j, j = 1, 2, \dots, J$ ,  $\epsilon$  is the error terms column vector, and  $\beta$  is the parameter column vector;  $X$  is the matrix of order  $J \times L$ , and it is defined as the design matrix. The architecture matrix has one column or explaining element and one row per observation per model parameter.

**3.2. Cumulants.** The estimated parameters of GLM are fed to the optimal detector block as input data, which includes feature extraction and prediction stages after it has been transformed and selected. The extraction block calculates statistics called cumulative characteristics consisting of moments. Let  $M_{pq}$  represents the moments of input data  $x(n)$  which are estimated parameters of GLM and is calculated using

$$M_{pq} = E\{x(n)^{p-q} x^*(n)^q\}, \quad (3)$$

where  $E\{\cdot\}$  is defined as the expected value function which calculates the expected value of a random variable, and  $p$  and  $q$  are the indexed power term to define cumulants.

$$E\{x\} = \int_{-\infty}^{\infty} xf(x)dx. \quad (4)$$

For the defined input data  $x(n)$ , cumulants  $\Psi_{ij}$  of the second and fourth order are as follows [34]:

$$\begin{aligned} \Psi_{20} &= E\{x^2(n)\}, \\ \Psi_{21} &= E\{|x(n)|^2\}, \\ \Psi_{40} &= M_{40} - 3M_{20}^2, \\ \Psi_{41} &= M_{40} - 3M_{20}M_{21}, \\ \Psi_{42} &= M_{42} - |M_{20}|^2 - 2M_{21}. \end{aligned} \quad (5)$$

**3.3. Likelihood Ratio Test.** The likelihood ratio test (LRT) is a test that is based on statistical analysis. It is a hypothesis test that tells how rightly two models can fit. This helps in choosing the best model between them [52]. The best model is chosen based on the likelihood function, means which maximizes the likelihood function.

Let  $Y = [y_1, y_2, y_3, \dots, y_k]$  represents the match scores for  $K$  individual matchers. The random variable  $y_k$  represents the  $k$ th matcher's match score, where  $k = 1, 2, \dots, K$ . Let's call the two classes  $M_0$  and  $M_1$ , where  $M_1$  denotes a true positive (genuine) class and  $M_0$  denotes a true negative class. The conditional joint densities of the  $k$  match scores assigned to the first and second classes, respectively, are  $p(Y|H_0)$  and  $p(Y|H_1)$ , where  $Y = [y_1, y_2, y_3, \dots, y_k]$ , and assume that the null hypothesis is  $H_0$  and the alternative hypothesis is  $H_1$ . Suppose the aim is to categorise the observed match score vector  $Y$  into one of two classes:  $M_0$  or  $M_1$ . The null hypothesis should be dismissed, and the alternative hypothesis should be accepted. The likelihood ratio between null and alternative hypotheses is evaluated and analyzed for a decision threshold  $\theta$ . The likelihood ratio test (LRT) is described as

$$LR(Y, H_0, H_1) = \frac{p(Y|H_0)}{p(Y|H_1)}, \quad (6)$$

where  $Y$  is the observed parameter,  $p(Y|H_0)$  is the likelihood function for the null hypothesis which is evaluated for  $Y$ , similarly  $p(Y|H_1)$  is the likelihood function for alternate hypothesis, and  $\theta$  is the decision threshold which decides the acceptance or rejection of the null hypothesis. For example, if  $LR(Y, H_0, H_1) \geq \theta$ , the null hypothesis is accepted; otherwise, the alternative hypothesis is accepted.

In this study, Kernel density estimation (KDE) [53, 60] is used along with LRT for the acceptance or rejection of the null/alternative hypothesis. The approximation of probability density functions (pdfs) for training data is done using KDE. This is a nonparametric approach and has benefited because there is no permanent structure in KDE. Moreover, during estimation in KDE, all the data points are included in the analysis.

**3.4. Proposed Algorithm.** The collected data are first transformed into a simple matrix having rows and columns. This transformation is done using GLM based on  $T$ -values which is explained in the next section. The  $\beta$  values found using the design matrix are used as initial features and are comprehensively discussed in the performance evaluation section. SPM is used for the initial analysis of the data.

The kernel density function which is also known as kernel density estimation (KDE) is used for estimated class value; that is, the best match of each class with the test vector is found. Features from 50 to 5000 have been used and transformed these features using cumulants to reduce the size of the features vector. As a result, a maximum of eight features has been formed instead of 50 or 5000. Now, it is easy for kernel density function in terms of computation to process and find the best match during the fusion of match scores. Moreover, in this study, one-to-one results are extended to multiclass classification or multiclass decoding. In this case, one out of five classes was found using the proposed method. The detail is discussed in the result section. For cross-validation, Monte Carlo sampling is used, and the accuracy is found based on the average of 100 trials for each class to make the result reliable. The pseudo-code of the proposed methodology is depicted in Algorithm 1.

## 4. Results and Discussion

In this section, a detailed performance evaluation of the proposed optimal predictor is presented. In the experiment design, we had different images which were divided into five different categories. These images were taken from [61], and the categories were also made based on the same study. The images of five different classes are human, animal images, images of the building, images of natural scenes, and images of fruits. The design matrix is a matrix of values that consists of explanatory variables of different conditions or categories. Since there are five different categories, one variable (regressor) is defined for every category in the design matrix. A total of five variables are defined in the design matrix for five different conditions, and furthermore, for baseline and realignment parameters, separate variables (regressors) are used in the design matrix. Six realignment parameters are used to remove noise also known as nuisance regressors. There are five conditions, six nuisance regressors and a baseline, so a total of 12 regressors are used in the design matrix. The activity difference is found between the condition and the baseline for each category separately. For example, the activity difference between humans and baseline is shown in Figure 2. The statistical analysis of the image (Figure 2) is mentioned in the corresponding Table 1. Table 1 explains the degree of freedom, full width at half maximum (FWHM), voxel size, the position of voxels,  $z$ -values, number of clusters, number of significant voxels in each cluster, and other information. In the figure, the red arrow ( $<$ ) shows the current location of the voxel (mostly the highest significant voxel), and the detail of this voxel is present in the first row (if highly significant) of Algorithm 1 with MNI coordinates, as also shown in red color. Since the significant voxels are important during analysis, so in the

- (i) **Data Transformation.**  
 (a) Estimated values of the voxels are found using the General linear model.  
 (b) The significant brain area is found and shown based on the highest  $T$ -value.
- (ii) **Data Selection**  
 (a) The significant features are selected based on  $p$ -values.  
 (b) Set several selected features having lower  $p$ -values.
- (iii) **Feature Extraction**  
 (a) Extract the HOCs from data  $x(n)$  using equation (5)  $[\Psi]_{i=1}^N$ , where  $N$  is number of selected features.
- (iv) **Likelihood Ratio Test**  
 (a) Do the estimation of the densities for different classes using the Kernel density estimator.  
 (b) The density-based score fusion is done using likelihood ratio-based score fusion  
 (c) The estimation of the score of the class is done using the likelihood ratio test using equation (6).

ALGORITHM 1: Optimal prediction algorithm.

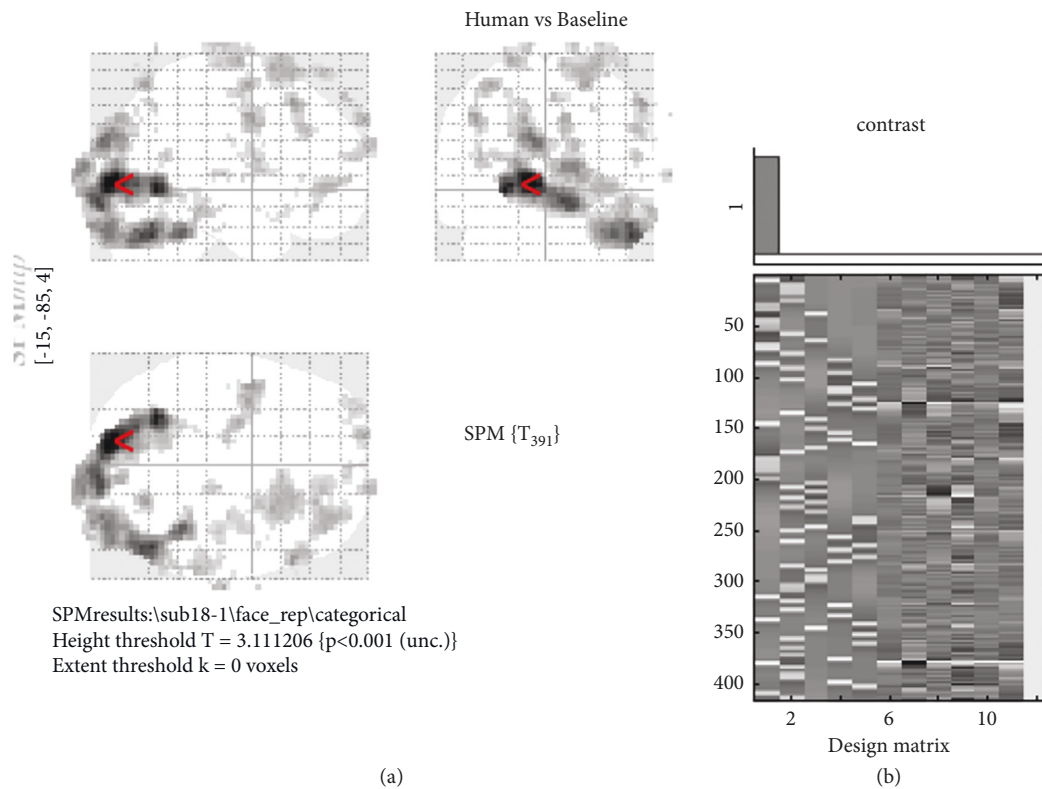


FIGURE 2: (a) (First-level analysis event-related design: human images vs baseline). (b) Design matrix which has 12 columns. In the design matrix, a number of variables are given horizontally, and trials are given vertically.

tables, the arrangement of voxels/clusters is shown in descending order means voxels/clusters with higher  $T$ -values are mentioned first. In Algorithm 1, the first three rows show the three most significant voxels of the most significant cluster while the first row shows the most significant voxels of the whole brain region for this task. In Figure 2, the brain activations are shown between humans and baseline while the behavior of the brain is the same for other sessions and participants. In the glass brain, it is observed that the activation area is the same for all the conditions (occipital region), but small differences can be seen among different categories. Generally, in the brain, there is always a small difference during the task concerning

baseline [57], so further statistical analysis is required to find that difference. In the first-level analysis, those significant voxels are found which shows the activity differences during the condition to baseline for all sessions of every participant. These voxels are just the  $\beta$ -value for each trial and each voxel which was then used as a feature [17] for further analysis. In the existing study [17], these  $\beta$ -values are directly used as features and given to the classifier, that is, SVM to find the accuracy. On the other hand, in the current study, these features are further refined using a different order of cumulants, and finally, instead of using SVM, a different statistical technique is used, that is, LRT. SVM is the most common, reliable, and widely used classifier in neuroscience

TABLE 1: Static table for human vs baseline shows the degree of freedom, FWHM, voxel size, the position of voxels,  $t$ -value, cluster, and other details.

Statistics: p-value adjusted for search volume													
Set level		Cluster level			Peak level						mm	mm	mm
$p$	$c$	$P_{FEW_{corr}}$	$q_{FDR_{corr}}$	$k_E$	$P_{uncorr}$	$P_{FEW_{corr}}$	$q_{FDR_{corr}}$	$T$	$Z$	$P_{uncorr}$			
0.002	45	0.000	0.000	2094	0.000	0.000	0.000	9.75	Inf	0.000	-15	-85	4
						0.000	0.000	9.67	Inf	0.000	-9	-91	1
						0.000	0.000	8.94	Inf	0.000	-27	-61	-2
		0.000	0.000	80	0.000	0.007	0.003	5.25	5.15	0.000	57	23	52
						0.328	0.047	4.44	4.38	0.000	54	32	34
		0.000	0.000	118	0.000	0.013	0.006	5.12	5.03	0.000	-45	-1	40
						0.029	0.010	4.96	4.88	0.000	-36	-4	61
						0.953	0.190	3.90	3.86	0.000	-45	-7	55
		0.005	0.001	39	0.000	0.026	0.010	4.98	4.9	0.000	24	-67	61
						1.000	0.676	3.35	3.32	0.000	30	-73	73
		0.004	0.001	40	0.000	0.037	0.012	4.91	4.83	0.000	-15	68	40
		0.000	0.000	248	0.000	0.071	0.020	4.77	4.70	0.000	27	11	70
						0.143	0.029	4.62	4.56	0.000	21	23	67
						0.221	0.037	4.52	4.46	0.000	9	17	73
		0.000	0.000	64	0.000	0.082	0.021	4.74	4.67	0.000	-24	-10	76
						0.628	0.081	4.24	4.19	0.000	-12	-19	85
						0.677	0.088	4.20	4.15	0.000	-18	-16	76
		0.205	0.021	16	0.008	0.082	0.021	4.74	4.67	0.000	75	-19	7
						1.000	0.777	3.27	3.25	0.001	69	-13	7
		0.000	0.000	123	0.000	0.087	0.021	4.73	4.66	0.000	33	47	1
						0.174	0.031	4.58	4.51	0.000	24	38	-11

Table shows three local maxima more than 0.8 mm apart. Height threshold:  $T = 3.11$ ,  $p = 0.001$  (1.000). Extent thresMd:  $k: 0$  voxels. Expected voxels per cluster,  $\langle k \rangle = 2.012$ . Expected number of clusters,  $\langle c \rangle \geq 28.41$ . FWEp: 4.846, FDRp: 4.427, FWEc: 27, FDRc: 14. Degree of freedom = [1.0, 3.91, 0]. FWHM: 7.7, 7.9, 7.6 mm mm mm; 2.6, 2.6, 2.5 (voxels). Volume: 1481436 = 54868 voxels = 2846.7 resels. Voxel size: 3.0 3.0 3.0 mm mm mm; (resel: 17.29 voxels).

studies, but we propose different classification technique which shows better results compared to SVM.

During the analysis, the raw fMRI data were cleaned and passed through GLM to find the  $\beta$ -values which were used as initial features since fMRI data have a lot of features so feature reduction or refinement is very common. These initial features were refined and passed through LRT to predict the right class. The accuracy between 5 different classes was found using the one-to-one decoding method [62] between every two classes separately. The same procedure was repeated for other existing methods as shown in Table 2 to see the performance of the proposed method. There were a total of 10 one-to-one decoding combinations due to five different classes. The data were randomly distributed between 90% of training data and 10% of test data using Monte Carlo cross-validation. The 10% test data evaluate the performance of the proposed method. One-vs-one decoding results are shown in Figure 3 for all 10 combinations along with the number of features. In Figure 3, it is mentioned that the best results exist between 400 – 500 features in most of the cases. Classification accuracy is found for every individual participant but after combining the voxels of all three sessions. Classification is done one-against-one with multiclass SVM between every condition and participant separately. An uncorrected  $p$ -value of 0.001 is used to find the significance of every category.

It is very common in existing studies to take initial features from fMRI data (either raw data or  $\beta$ -value) and used SVM to find the prediction accuracy [17, 19, 63, 64].

In this study, we have proposed a different method that is more refined in terms of significant features and predictors. That's why to check the performance of the proposed system, the results are compared with various existing methods. The performance is checked based on accuracy and time which is one of the important factors. The proposed method is compared with various other methods mentioned in Table 2 and is also shown graphically in Figure 4. In linear discriminant analysis (LDA),  $\beta$ -values were taken as features as directly given to the classifier [17]. The features were not refined like the proposed method. Similarly, for the LibSVM case, same  $\beta$ -values were taken as features, and LibSVM was used as a classifier [65]. In the case of LRBSF, same features were used for the further process, but instead of any existing classifier, LRT was performed to find the best match of the predictor. In the existing study, LRT is used with refined features, that is, features found using cumulants.

For multiclass discrimination, the above-mentioned method (2 class likelihood ratio test) is extended to multiclass. Instead of choosing the best score match between two classes, we have chosen the match score with the highest value between the match score of the test vector and the class type. The result of multiclass decoding is shown in Figure 5 which indicates the accuracy against the number of features for all five classes simultaneously. The result is approximately 35% which is quite good as the chance in multidecoding is quite low, especially in 5 classes. Multiclass likelihood test is also used using LRBSF.

TABLE 2: Accuracy comparison of various methods using the fMRI data set.

Condition	LDA (%)	Deep learning (%)	Lib SVM (%)	LRBSF (%)	Proposed method (%)
Human vs animal	58.10	60.60	59.50	61.15	66.98
Human vs building	62.40	64.14	63.27	64.92	70.43
Human vs natural scenes	65.95	67.94	67.60	68.37	71.84
Human vs fruit	62.84	67.20	64.13	67.79	69.96
Animal vs building	65.25	68.80	66.23	69.42	71.71
Animal vs natural scenes	68.84	70.95	67.49	71.30	73.66
Animal vs fruit	66.12	67.17	67.74	67.73	68.88
Building vs natural scenes	58.54	60.10	57.68	59.47	67.72
Building vs fruit	62.29	64.25	63.93	64.62	70.10
Natural scenes vs fruit	62.47	64.84	63.92	65.52	69.82
Average	63.25	65.60	64.10	66.10	70.11

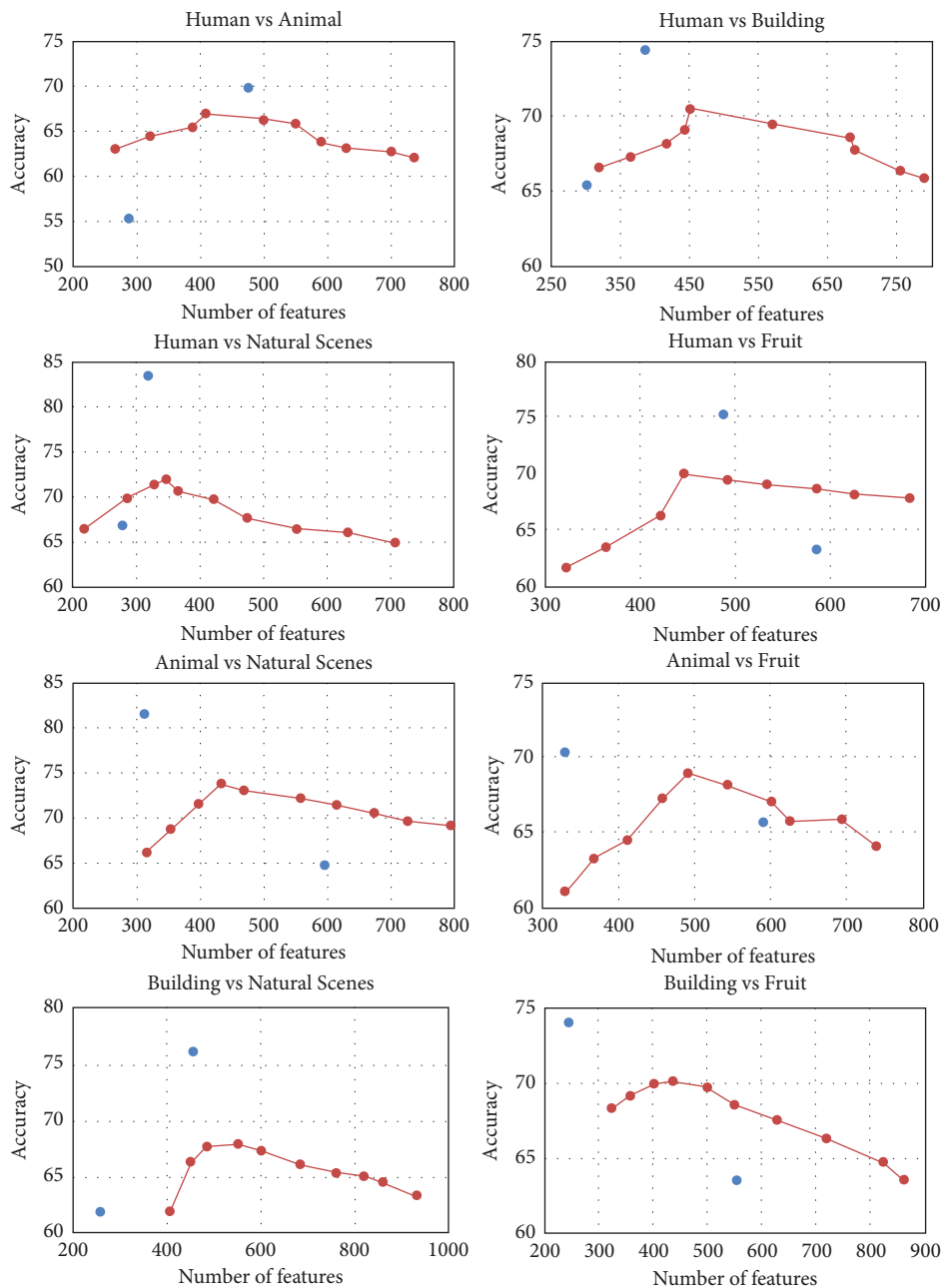


FIGURE 3: Average prediction accuracy of various combinations against the number of features. One-versus-one prediction for five different conditions which made 10 different combinations.

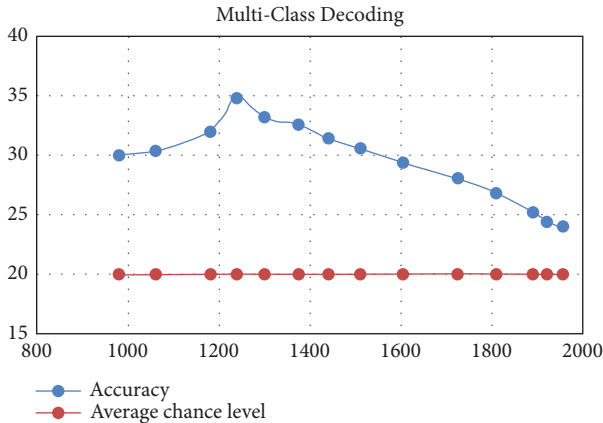


FIGURE 4: Comparison of various methods with the proposed method.

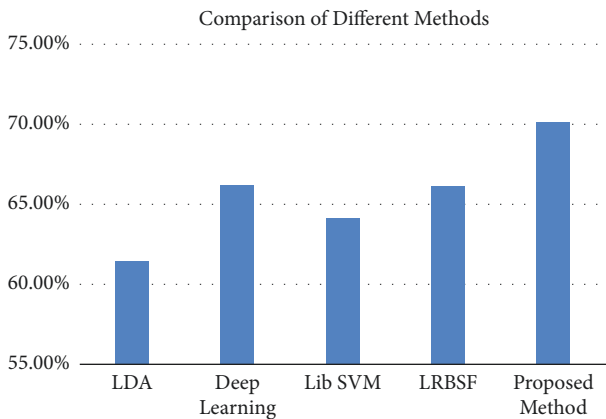


FIGURE 5: Multiclass decoding accuracy of five different classes. The chance level for five different classes is 20% while the best result is 35% according to the proposed method.

A comprehensive analysis of brain response and behavior is done, and the new features are used to improve the accuracy and response time. This study will help in various ways as it is contributing in many ways. The likelihood ratio test (LRT) is applied with cumulant features along with MVPA, and moreover, different classifiers are used to compare the results with LRT. This study gives the basic knowledge as well as focuses on advanced statistical and mathematical methods like LRT and the role of different levels of cumulants for the fMRI data set.

In short, this study is an addition to the existing methods proposed for fMRI data analysis. Since hardware limitations are not easy to overcome, new statistical and mathematical methods are always helpful in the analysis of neuronal data including fMRI, EEG, and MEG data.

## 5. Conclusion

In this study, a hybrid method is proposed which is an extension of various existing methods such as basic  $\beta$ -values or raw fMRI data are used as features for further analysis in most of the existing studies [17, 64]. In this study, these

$\beta$ -values are further polished to make the features more significant. Similarly, instead of using SVM which is the most common and widely technique of classification, a different statistical technique is used in neuroscience, that is, LRBSF. This technique is widely used in other applications [52, 66, 67] but has never been used in the analysis of fMRI data before. This LRBSF technique was initially introduced for fMRI data in 2017 by our research group [53]. In other words, this study is an extension and improvement of the previous study in terms of accuracy and response time. In the previous study, the response time was quite high which is improved in this study along with the accuracy. In short, this study includes numerous statistical and mathematical methods during the extraction of features. These new features are combined with the existing method, that is, LRBSF, which provides a novel hybrid method for the analysis of brain data acquired using fMRI. The new features which are introduced in the proposed algorithm are different levels of cumulants, and these features improved the overall accuracy from 66% to 70% while less response time is used than the previous one [53].

In the future, this algorithm can be applied to other modalities such as EEG and MEG. Moreover, other statistical methods such as the Gaussian mixture model and sequential Markov chain methods [27] and further levels of cumulants can be used with the existing methods to see the effect of additional features. This will be a new challenge for the researchers to work on various statistical techniques along with different modalities. The purpose is to achieve sufficient accuracy so that the neuroscience applications can widely be used especially BCI applications which can easily work with EEG data. By achieving acceptable accuracy, the fear among people can be reduced and actively build public support for neuroscience research. The goals should be set so that the public can recognize this field and share.

## Data Availability

The data were collected at the Hospital Universiti Sains Malaysia (HUSM). The details of the data are given in the manuscript. The processed data can be provided on demand. In our study, we have normal subjects, not patients, and we only examine the effect (for example, provider knowledge or attitudes). Therefore, the International Committee of Medical Journal Editors (ICMJE) on trial registrations is not required.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## References

- [1] V. Lekka, "Mapping the relations between history and history of science: the case of the history of psychiatry," *Rethinking History*, vol. 21, no. 4, pp. 606–617, 2017.
- [2] Y. Frégnac and G. Laurent, "Neuroscience: where is the brain in the human brain project?" *Nature News*, vol. 513, no. 7516, p. 27, 2014.

- [3] V. Hughes, "Science in court: head case," *Nature*, vol. 464, no. 7287, pp. 340–342, 2010.
- [4] P. M. Matthews, G. D. Honey, and E. T. Bullmore, "Applications of fmri in translational medicine and clinical practice," *Nature Reviews Neuroscience*, vol. 7, no. 9, pp. 732–744, 2006.
- [5] M. S. Walker, A. M. Stamper, D. E. Nathan, and G. Riedy, "Art therapy and underlying fmri brain patterns in military tbi: a case series," *International Journal of Art Therapy*, vol. 23, no. 4, pp. 180–187, 2018.
- [6] M. S. Sherwood, E. E. Diller, E. Ey, S. Ganapathy, J. T. Nelson, and J. G. Parker, "A protocol for the administration of real-time fmri neurofeedback training," *Journal of Visualized Experiments: Journal of Visualized Experiments*, vol. 126, 2017.
- [7] G. Van Doorn and B. Miloyan, "The pepsi paradox: a review," *Food Quality and Preference*, vol. 65, pp. 194–197, 2018.
- [8] S. M. McClure, J. Li, D. Tomlin, K. S. Cypert, L. M. Montague, and P. Montague, "Neural correlates of behavioral preference for culturally familiar drinks," *Neuron*, vol. 44, no. 2, pp. 379–387, 2004.
- [9] R. Jain, N. Jain, A. Aggarwal, and D. J. Hemanth, "Convolutional neural network based alzheimer's disease classification from magnetic resonance brain images," *Cognitive Systems Research*, vol. 57, pp. 147–159, 2019.
- [10] M. D'Esposito, A. S. Kayser, and A. J. Chen, "Functional mri: cognitive neuroscience applications," in *Functional Neuro-radiology* Springer, Berlin, Germany, 2011.
- [11] M. Carbó-Carreté, C. Cañete-Massé, M. Peró-Cebollero, and J. Guardia-Olmos, "Using fmri to assess brain activity in people with down syndrome: a systematic review," *Frontiers in Human Neuroscience*, vol. 14, p. 147, 2020.
- [12] B. Gutkin, D. Pinto, and B. Ermentrout, "Mathematical neuroscience: from neurons to circuits to systems," *Journal of Physiology Paris*, vol. 97, no. 2-3, pp. 209–219, 2003.
- [13] T. Trappenberg, *Fundamentals of Computational Neuroscience*, OUP Oxford, Oxford, UK, 2009.
- [14] A. U. Rehman, R. A. Naqvi, A. Rehman, A. Paul, M. T. Sadiq, and D. Hussain, "A trustworthy snot aware mechanism as an enabler for citizen services in smart cities," *Electronics*, vol. 9, no. 6, 2020.
- [15] A. U. Rehman, A. Jiang, and A. Rehman, "Weighted based trustworthiness ranking in social internet of things by using soft set theory," in *Proceedings of the IEEE 5th International Conference on Computer and Communications (ICCC)*, pp. 1644–1648, IEEE, Chengdu, China, December 2019.
- [16] C.-A. Chou, K. Kampa, S. H. Mehta, R. F. Tugaraza, W. A. Chaovalitwongse, and T. J. Grabowski, "Voxel selection framework in multi-voxel pattern analysis of fmri data for prediction of neural response to visual stimuli," *IEEE Transactions on Medical Imaging*, vol. 33, no. 4, pp. 925–934, 2014.
- [17] A. Gardumi, D. Ivanov, L. Hausfeld, G. Valente, E. Formisano, and K. Uludag, "The effect of spatial resolution on decoding accuracy in fmri multivariate pattern analysis," *NeuroImage*, vol. 132, pp. 32–42, 2016.
- [18] M. J. Rosa and B. Seymour, "Decoding the matrix: benefits and limitations of applying machine learning algorithms to pain neuroimaging," *Pain*, vol. 155, no. 5, pp. 864–867, 2014.
- [19] T. Horikawa, M. Tamaki, Y. Miyawaki, and Y. Kamitani, "Neural decoding of visual imagery during sleep," *Science*, vol. 340, no. 6132, pp. 639–642, 2013.
- [20] R. Zafar, N. Kamel, M. Naufal et al., "A study of decoding human brain activities from simultaneous data of eeg and fmri using mvpa," *Australasian Physical & Engineering Sciences in Medicine*, vol. 41, no. 3, pp. 633–645, 2018.
- [21] R. A. Poldrack, "Region of interest analysis for fmri," *Social Cognitive and Affective Neuroscience*, vol. 2, no. 1, pp. 67–70, 2007.
- [22] J. V. Haxby, M. I. Gobbini, M. L. Furey, A. Ishai, J. L. Schouten, and P. Pietrini, "Distributed and overlapping representations of faces and objects in ventral temporal cortex," *Science*, vol. 293, no. 5539, pp. 2425–2430, 2001.
- [23] M. Cauchoix, G. Barragan-Jason, T. Serre, and E. J. Barbeau, "The neural dynamics of face detection in the wild revealed by mvpa," *Journal of Neuroscience*, vol. 34, no. 3, pp. 846–854, 2014.
- [24] R. Zafar, M. N. Malik, and H. Hayat, "Decoding brain patterns for colored and grayscale images using multivariate pattern analysis," *KSII Transactions on Internet and Information Systems (TIIS)*, vol. 14, no. 4, pp. 1543–1561, 2020.
- [25] S. M. Hayes, L. Ryan, D. M. Schnyer, and L. Nadel, "An fmri study of episodic memory: retrieval of object, spatial, and temporal information," *Behavioral Neuroscience*, vol. 118, no. 5, pp. 885–896, 2004.
- [26] R. Zafar, S. C. Dass, and A. S. Malik, "Electroencephalogram-based decoding cognitive states using convolutional neural network and likelihood ratio based score fusion," *PLoS One*, vol. 12, no. 5, Article ID e0178410, 2017.
- [27] P. Agrawal, *A Probe into Decoding Brain Activity Using Fmri and Meg*, 2011.
- [28] S. A. Huettel, A. W. Song, and G. McCarthy, *Functional Magnetic Resonance Imaging*, Vol. 1, Sinauer Associates, , Sunderland, MA, USA, 2004.
- [29] T. Grootswagers, S. G. Wardle, and T. A. Carlson, "Decoding dynamic brain patterns from evoked responses: a tutorial on multivariate pattern analysis applied to time series neuro-imaging data," *Journal of Cognitive Neuroscience*, vol. 29, no. 4, pp. 677–697, 2017.
- [30] A. U. Rehman, A. Jiang, A. Rehman, A. Paul, S. din, and M. T. Sadiq, "Identification and role of opinion leaders in information diffusion for online discussion network," *Journal of Ambient Intelligence and Humanized Computing*, vol. 1–13, 2020.
- [31] A. G. Huth, T. Lee, S. Nishimoto, N. Y. Bilenko, A. T. Vu, and J. L. Gallant, "Decoding the semantic content of natural movies from human brain activity," *Frontiers in Systems Neuroscience*, vol. 10, p. 81, 2016.
- [32] W. U. Khan, T. N. Nguyen, F. Jameel, and M. A. Jamshed, "Learning-based resource allocation for backscatter-aided vehicular networks," *IEEE Transactions on Intelligent Transportation Systems*, 2021.
- [33] A. Mahmood, Y. Hong, M. K. Ehsan, and S. Mumtaz, "Optimal resource allocation and task segmentation in iot enabled mobile edge cloud," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 12, Article ID 13294, 2021.
- [34] W. U. Khan, A. Ihsan, T. N. Nguyen, M. A. Javed, and Z. Ali, "Noma-enabled backscatter communications for green transportation in automotive-industry 5.0," *IEEE Transactions on Industrial Informatics*, 2022.
- [35] A. Mahmood, A. Ahmed, M. Naem, and Y. Hong, "Partial offloading in energy harvested mobile edge computing: a direct search approach," *IEEE Access*, vol. 8, Article ID 36757, 2020.
- [36] A. U. Rehman, R. Tariq, A. Rehman, and A. Paul, "Collapse of online social networks: structural evaluation, open challenges, and proposed solutions," in *Proceedings of the IEEE Globecom Workshops (GC Wkshps)*, vol. 1–6, IEEE, Taipei, Taiwan, December 2020.

- [37] J. R. McIntosh, J. Yao, L. Hong, J. Faller, and P. Sajda, "Ballistocardiogram artifact reduction in simultaneous eeg-fmri using deep learning," *IEEE Transactions on Biomedical Engineering*, vol. 68, no. 1, pp. 78–89, 2021.
- [38] J. Jorge, F. Grouiller, Ö. Ipek et al., "Simultaneous eeg–fmri at ultra-high field: artifact prevention and safety assessment," *NeuroImage*, vol. 105, pp. 132–144, 2015.
- [39] R. J. Huster, S. Debener, T. Eichele, and C. S. Herrmann, "Methods for simultaneous eeg-fmri: an introductory review," *Journal of Neuroscience*, vol. 32, no. 18, pp. 6053–6060, 2012.
- [40] M. A. Nielsen, *Neural Networks and Deep Learning*, Vol. 25, Determination press, San Francisco, CA, USA, 2015.
- [41] A. X. Stewart, A. Nuthmann, and G. Sanguinetti, "Single-trial classification of eeg in a visual object task using ica and machine learning," *Journal of Neuroscience Methods*, vol. 228, pp. 1–14, 2014.
- [42] A. Subasi and M. Ismail Gursoy, "Eeg signal classification using pca, ica, lda and support vector machines," *Expert Systems with Applications*, vol. 37, no. 12, pp. 8659–8666, 2010.
- [43] R. Zafar, M. J. ur Rehman, and H. Dilpazir, "Power spectrum of brain for complex task using electroencephalogram," *Technology*, vol. 4, no. 2, pp. 383–391, 2022.
- [44] E. Derya Übeyli, "Statistics over features: eeg signals analysis," *Computers in Biology and Medicine*, vol. 39, no. 8, pp. 733–741, 2009.
- [45] M. J. U. Rehman, R. Zafar, H. Dilpazir, M. F. Sohail, M. A. Khan, and R. Jahan, "Embedded estimation sequential bayes parameter inference for the ricker dynamical system," *Journal of Sensors*, vol. 2022, pp. 1–8, 2022.
- [46] M. J. ur Rehman, S. C. Dass, and V. S. Asirvadam, "A weighted likelihood criteria for learning importance densities in particle filtering," *EURASIP Journal on Applied Signal Processing*, vol. 2018, no. 1, pp. 1–19, 2018.
- [47] M. K. Ehsan and D. Dahlhaus, "Statistical modeling of ism data traffic in indoor environments for cognitive radio systems," in *Proceedings of the Third International Conference on Digital Information, Networking, and Wireless Communications (DINWC)*, vol. 88–93, IEEE, Moscow, Russia, February 2015.
- [48] A. Mahmood, A. Ahmed, M. Naeem, M. R. Amirzada, and A. Al-Dweik, "Weighted utility aware computational overhead minimization of wireless power mobile edge cloud," *Computer Communications*, vol. 190, pp. 178–189, 2022.
- [49] M. Javvad ur Rehman, S. C. Dass, and V. S. Asirvadam, "An augmented sequential mcmc procedure for particle based learning in dynamical systems," *Signal Processing*, vol. 160, pp. 32–44, 2019.
- [50] M. J. ur Rehman, S. C. Dass, and V. S. Asirvadam, "A bayesian parameter learning procedure for nonlinear dynamical systems via the ensemble kalman filter," in *Proceedings of the IEEE 14th International Colloquium on Signal Processing & Its Applications (CSPA)*, vol. 161–166, IEEE, Penang, Malaysia, March 2018.
- [51] M. J. ur Rehman, S. C. Dass, V. S. Asirvadam, and A. Adly, "Parameter estimation for nonlinear disease dynamical system using particle filter," in *Proceedings of the International Conference on Computer, Control, Informatics and its Applications (IC3INA)*, vol. 143–147, IEEE, Bandung, Indonesia, October 2015.
- [52] K. Nandakumar, Y. Chen, S. C. Dass, and A. Jain, "Likelihood ratio-based biometric score fusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 2, pp. 342–347, 2008.
- [53] R. Zafar, S. C. Dass, A. S. Malik et al., "Prediction of human brain activity using likelihood ratio based score fusion," *IEEE Access*, vol. 5, Article ID 13010, 2017.
- [54] A. Hussain, M. F. Sohail, S. Alam, S. A. Ghauri, and I. M. Qureshi, "Classification of m-qam and m-psk signals using genetic programming (gp)," *Neural Computing & Applications*, vol. 31, no. 10, pp. 6141–6149, 2019.
- [55] M. Sarfraz, S. Alam, S. A. Ghauri et al., "Random graph-based m-qam classification for mimo systems," *Wireless Communications and Mobile Computing*, pp. 1–10, 2022.
- [56] W. D. Hairston and J. A. Maldjian, "An adaptive staircase procedure for the e-prime programming environment," *Computer Methods and Programs in Biomedicine*, vol. 93, no. 1, pp. 104–108, 2009.
- [57] K. J. Friston, P. Jezzard, and R. Turner, "Analysis of functional mri time-series," *Human Brain Mapping*, vol. 1, no. 2, pp. 153–171, 1994.
- [58] J. Ashburner, K. Friston, and W. Penny, Chapter 37—the general linear model, 2004.
- [59] I. Winkler, S. Brandl, F. Horn, E. Waldburger, C. Allefeld, and M. Tangermann, "Robust artifactual independent component classification for bci practitioners," *Journal of Neural Engineering*, vol. 11, no. 3, Article ID 035013, 2014.
- [60] S. J. Sheather and M. C. Jones, "A reliable data-based bandwidth selection method for kernel density estimation," *Journal of the Royal Statistical Society: Series B*, vol. 53, no. 3, pp. 683–690, 1991.
- [61] K. N. Kay, T. Naselaris, R. J. Prenger, and J. L. Gallant, "Identifying natural images from human brain activity," *Nature*, vol. 452, no. 7185, pp. 352–355, 2008.
- [62] S. Cruzet, N. Busch, and K. Ohla, "Taste quality decoding parallels taste sensations," *Current Biology*, vol. 25, no. 7, pp. 890–896, 2015.
- [63] P. K. Douglas, E. Lau, A. Anderson et al., "Single trial decoding of belief decision making from eeg and fmri data using independent components features," *Frontiers in Human Neuroscience*, vol. 7, 2013.
- [64] J. A. Mumford, B. O. Turner, F. G. Ashby, and R. A. Poldrack, "Deconvolving bold activation in event-related designs for multivoxel pattern classification analyses," *NeuroImage*, vol. 59, no. 3, pp. 2636–2643, 2012.
- [65] C.-C. Chang and C.-J. Lin, "Libsvm: a library for support vector machines," *ACM Trans. Intell. Syst. Technol.* vol. 2, no. 3, pp. 1–27, 2011.
- [66] L. Nanni, A. Lumini, and S. Brahmam, "Likelihood ratio based features for a trained biometric score fusion," *Expert Systems with Applications*, vol. 38, no. 1, pp. 58–63, 2011.
- [67] M. W. Mitchell, M. G. Genton, and M. L. Gumpertz, "A likelihood ratio test for separability of covariances," *Journal of Multivariate Analysis*, vol. 97, no. 5, pp. 1025–1043, 2006.